

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 15:44:37 ; Search time 401.589 Seconds
(without alignments)
1449.381 Million cell updates/sec

Title: US-10-052-092-15

Perfect score: 20

Sequence: 1 caagcgccagagatgatg 20

Scoring table: IDENTITY_NTC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:*

1: gb_ha:*

2: gb_htg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_in:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_man:*

37: em_htg_vtl:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 20 | 100.0 | 26 | E13443 | E13443 PCR primer |
| 2 | 20 | 100.0 | 144 | S80316 | S80316 oestrogen r |
| 3 | 20 | 100.0 | 392 | AF120105 | AF120105 Homo sapi |
| 4 | 20 | 100.0 | 521 | AX004781 | AX004781 Sequence |
| 5 | 20 | 100.0 | 539 | AX004793 | AX004793 Sequence |
| 6 | 20 | 100.0 | 566 | AX004790 | AX004790 Sequence |
| 7 | 20 | 100.0 | 1223 | AX066401 | AX066401 Sequence |
| 8 | 20 | 100.0 | 1237 | HSESR13 | AF123496 Homo sapi |
| 9 | 20 | 100.0 | 1374 | AX066402 | AX066402 Sequence |
| 10 | 20 | 100.0 | 1788 | SSSTRRC | Z37167 S. scrofa mr |
| 11 | 20 | 100.0 | 1788 | AX111751 | AX111751 Sequence |
| 12 | 20 | 100.0 | 1966 | OAESTRRC | Z49257 O. aries MRN |
| 13 | 20 | 100.0 | 2092 | AX474709 | AX474709 Sequence |
| 14 | 20 | 100.0 | 2092 | I08538 | I08538 Sequence 1 |
| 15 | 20 | 100.0 | 2092 | HUMRMCF | M12674 Human estro |
| 16 | 20 | 100.0 | 2106 | BSU47678 | U47678 Human 80 KD |
| 17 | 20 | 100.0 | 2322 | AA2099 | AA2099 Sequence 1 |
| 18 | 20 | 100.0 | 3746 | AF124093 | AF124093 Equus cab |
| 19 | 20 | 100.0 | 4963 | AR029418 | AR029418 Sequence |
| 20 | 20 | 100.0 | 4963 | I15368 | I15368 Sequence 16 |
| 21 | 20 | 100.0 | 4963 | I56760 | I56760 Sequence 9 |
| 22 | 20 | 100.0 | 6553 | AY033393 | AY033393 Ovis arie |
| 23 | 20 | 100.0 | 6450 | ARI53585 | ARI53585 Sequence |
| 24 | 20 | 100.0 | 6450 | HSESR | X03635 Human mRNA |
| 25 | 20 | 100.0 | 6610 | AF061181 | AF061181 Mammalian |
| 26 | 20 | 100.0 | 6639 | AX128351 | AX128351 Sequence |
| 27 | 20 | 100.0 | 6639 | AX128347 | AX128347 Sequence |
| 28 | 20 | 100.0 | 6695 | AX128353 | AX128353 Sequence |
| 29 | 20 | 100.0 | 6695 | AX128354 | AX128354 Sequence |
| 30 | 20 | 100.0 | 6801 | AX128355 | AX128355 Sequence |
| 31 | 20 | 100.0 | 6818 | AX128346 | AX128346 Sequence |
| 32 | 20 | 100.0 | 6833 | AX128349 | AX128349 Sequence |
| 33 | 20 | 100.0 | 6900 | AX128341 | AX128341 Sequence |
| 34 | 20 | 100.0 | 6956 | AX128348 | AX128348 Sequence |
| 35 | 20 | 100.0 | 7038 | AX128342 | AX128342 Sequence |
| 36 | 20 | 100.0 | 139687 | AL590993 | AL590993 Human DNA |
| 37 | 20 | 100.0 | 165237 | AX232505 | AX232505 Sequence |
| 38 | 20 | 100.0 | 349980 | AX232503 | AX232503 Sequence |
| 39 | 20 | 100.0 | 349980 | AX453703 | AX453703 Sequence |
| 40 | 18.4 | 92.0 | 748 | CPA313195 | AT313195 Canis fam |
| 41 | 18 | 90.0 | 200000 | AC007355 | AC007355 Homo sapi |
| 42 | 17.4 | 90.0 | 243358 | AC110814 | AC110814 Homo sapi |
| 43 | 17.4 | 87.0 | 96599 | AC125306 | AC125306 Rattus no |
| 44 | 17.4 | 87.0 | 146067 | OSJN00150 | AL662948 Oryza sat |
| 45 | 17.4 | 87.0 | 160889 | AC109098 | AC109098 Rattus no |

ALIGNMENTS

RESULT 1
E13443/c
LOCUS E13443 26 bp DNA linear PAT 27-APR-1998
DEFINITION PCR primer for detecting mRNA which encode human estrogen receptor.
ACCESSION E13443
VERSION E13443.1 GI:3252248
KEYWORDS JP 1997187299-A/5.
SOURCE JP 1997187299-A/5.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 26)
AUTHORS KIMOTO Y.
TITLE PRIMER FOR PCR
JOURNAL Patent: JP 1997187299-A 5 22-JUL-1997;

COMMENT NIPPON BIO SERAPII KK
OS None
OC Artificial sequences.
PN JP 1997187299-A/5
PD 22-JUL-1997
PF 05-JAN-1996 JP 1996027222
PI KIMONO YASUHIKO
PC C1201/68,C07H21/04,C12M15/09;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: Yes;
FH Key Location/Qualifiers
FT source 1..26
FT /organism='Artificial sequences' FT
misc_feature 1..26
FT /note='PCR primer E-7'.

FEATURES
source Location/Qualifiers
1..26
/organism='unidentified'
/db_xref='taxon:32644'

BASE COUNT 2 a 8 c 5 g 11 t

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 20 CAAGCGCCAGAGAGATGATG 1
|||||

RESULT 2
S80316 144 bp mRNA linear PRI 02-APR-1996
LOCUS oestrogen receptor [alternatively spliced, exons 4 and 7] [human,
DEFINITION endometrial liver, mRNA Partial Mutant, 144 nt].
ACCESSION S80316
VERSION S80316.1 GI:1246086
KEYWORDS
SOURCE Homo sapiens endometrial liver.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 144)
AUTHORS Daifada,A.A. and Dowsett,M.
TITLE Tissue-dependent expression of a novel splice variant of the human
oestrogen receptor
J. Steroid Biochem. Mol. Biol. 55 (3-4), 413-421 (1995)
MEDLINE 96132694
PUBMED 8541239
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq.173922] from the original journal article.
This sequence comes from Fig. 3.

FEATURES
source Location/Qualifiers
1..144
/organism='Homo sapiens'
/db_xref='taxon:9606'
1..144
/partial
/gene='oestrogen receptor, delta 4/7 ER'
1..135
/partial
/gene='oestrogen receptor, delta 4/7 ER'
/note='this sequence comes from Fig. 3; delta 4/7 ER'
/codon_start=1
/product='oestrogen receptor'
/protein_id='PAB35900.1'
/db_xref='GI:1246087'
/translation='KHKRODDGEGRPDPDQGRPFAAAPAAGPAPHPHPQAH'

CDS
/partial
/gene='oestrogen receptor, delta 4/7 ER'
1..135
/partial
/gene='oestrogen receptor, delta 4/7 ER'
/note='this sequence comes from Fig. 3; delta 4/7 ER'
/codon_start=1
/product='oestrogen receptor'
/protein_id='PAB35900.1'
/db_xref='GI:1246087'
/translation='KHKRODDGEGRPDPDQGRPFAAAPAAGPAPHPHPQAH'

BASE COUNT 38 a 48 c 40 g 18 t

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 144;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 6 CAAGCGCCAGAGAGATGATG 25
|||||

RESULT 3
AF120105 392 bp mRNA linear PRI 06-APR-1999
LOCUS Homo sapiens alternatively-spliced estrogen receptor alpha mRNA,
DEFINITION partial cds.
ACCESSION AF120105
VERSION AF120105.1 GI:4567039
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS Campbell-Thompson,M.L.
TITLE Direct Submision
JOURNAL Submitted (13-JAN-1999) Medicine, University of Florida, Box 100214
Gastroenterology, Gainesville, FL 32667, USA

FEATURES
source Location/Qualifiers
1..392
/organism='Homo sapiens'
/db_xref='taxon:9606'
/map='6625.1'
/sex='male'
/cell_type='colon adenocarcinoma'
1..343
/note='steroid hormone receptor; missing exon 5'
/codon_start=2
/product='alternatively-spliced estrogen receptor alpha'
/protein_id='P4D23565.1'
/db_xref='GI:4567040'
/translation='RRGRMLKHKRQDDGEGREVGASGDMNANLWPSPLMKRSK
KNSLSTLADQWVSALDAEPPIIYSEDPTRFSPASMMGILTNLADRELIVMIW
AKRYGTFRENV'
25
/note='compared to wild type sequence deposited in GenBank
Accession Number X03633'
/replace='a'
92 c 124 g 79 t

BASE COUNT 97 a 92 c 124 g 79 t

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 392;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 28 CAAGCGCCAGAGAGATGATG 47
|||||

RESULT 4
AX004781 521 bp DNA linear PAT 24-AUG-2000
LOCUS AX004781
DEFINITION Sequence 1 from Patent WO9311785.
ACCESSION AX004781
VERSION AX004781.1 GI:9928194
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 521)

AUTHORS Mcgregor,D.
TITLE Chimeric binding peptide library screening method
JOURNAL Patent: WO 9911785-A 1 11-MAR-1999;
MCGREGOR DUNCAN (GB); ROWETT RESEARCH SERVICES LIMIT (GB)
FEATURES
SOURCE 1. 521
/organism="unidentified"
/db_xref="taxon:32644"
/note="Recombinant human oestrogen"
CDS 41. 124
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC04912.1"
/db_xref="GI:9928195"
/translation="MKTLPTAAAGLLLAQPMACVQLQ"
BASE COUNT 153 a 109 c 155 g 104 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 521;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
|||||
DB 427 CAAGCGCCAGAGAGATGATG 446
RESULT 5
AX004793 539 bp DNA linear PAT 24-AUG-2000
LOCUS Sequence 13 from Patent WO9911785.
DEFINITION AX004793
ACCESSION AX004793
VERSION AX004793.1 GI:9928204
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 539)
AUTHORS Mcgregor,D.
TITLE Chimeric binding peptide library screening method
JOURNAL Patent: WO 9911785-A 13 11-MAR-1999;
MCGREGOR DUNCAN (GB); ROWETT RESEARCH SERVICES LIMIT (GB)
FEATURES
SOURCE 1. 539
/organism="unidentified"
/db_xref="taxon:32644"
/note="Recombinant human oestrogen"
CDS 41. 481
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC04917.1"
/db_xref="GI:9928205"
/translation="MKTLPTAAAGLLLAQPMAMAKESAKETRYCANCNDYASGIH
YGVWSCGCKAFPKRSIQGHNDYCPAINOCTIDNRKSCQACRLKCYEVMKNG
IKKDRGGRMLKHKRDDEGRGEGSGGSSAQPALLDIAAA"
BASE COUNT 148 a 118 c 165 g 108 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 539;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
|||||
DB 385 CAAGCGCCAGAGAGATGATG 404
RESULT 6
AX004790 566 bp DNA linear PAT 24-AUG-2000
LOCUS Sequence 10 from Patent WO9911785.
DEFINITION AX004790
ACCESSION AX004790
VERSION AX004790.1 GI:9928202

KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 566)
AUTHORS Mcgregor,D.
TITLE Chimeric binding peptide library screening method
JOURNAL Patent: WO 9911785-A 10 11-MAR-1999;
MCGREGOR DUNCAN (GB); ROWETT RESEARCH SERVICES LIMIT (GB)
FEATURES
SOURCE 1. 566
/organism="unidentified"
/db_xref="taxon:32644"
/note="Recombinant human oestrogen"
CDS 41. 124
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC04916.1"
/db_xref="GI:9928203"
/translation="MKTLPTAAAGLLLAQPMAMAVQLQ"
BASE COUNT 164 a 118 c 169 g 115 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 566;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
|||||
DB 427 CAAGCGCCAGAGAGATGATG 446
RESULT 7
AX066401 1223 bp DNA linear PAT 24-JAN-2001
LOCUS Sequence 1 from Patent WO0100823.
DEFINITION AX066401
ACCESSION AX066401
VERSION AX066401.1 GI:12544111
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1223)
AUTHORS Gannon,F., Denger,S. and Flouriot,G.
TITLE Novel isoforms of the human estrogen receptor -g(a)
JOURNAL Patent: WO 0100823-A 1 04-JAN-2001;
EUROPEAN MOLECULAR BIOLOGY LABORATORY (DE)
FEATURES
SOURCE 1. 1223
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 303 a 316 c 341 g 263 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1223;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
|||||
DB 236 CAAGCGCCAGAGAGATGATG 255
RESULT 8
HSESRI3 1237 bp DNA linear PRI 07-APR-2000
LOCUS Homo sapiens estrogen receptor alpha (ESR1) gene, exon 4.
DEFINITION HSESRI3
ACCESSION AF123496
VERSION AF123496.1 GI:5821720
KEYWORDS
SEGMENT 3 of 7
SOURCE Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1237)
AUTHORS Schbert, E.L., Lee, M.K., Newman, B. and King, M.C.
TITLE Single nucleotide polymorphisms (SNPs) in the estrogen receptor
gene and breast cancer susceptibility
JOURNAL J. Steroid Biochem. Mol. Biol. 71 (1-2), 21-27 (1999)
MEDLINE 20084372
PUBMED 10619354
REFERENCE 2 (bases 1 to 1237)
AUTHORS Schbert, E.L., Lee, M.K. and King, M.C.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1999) Medical Genetics, University of Washington,
1959 NE Pacific Street, Seattle, WA 98195-7720, USA
FEATURES
Source
1. 1237
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q25"
/gene="ESR1"
/number="4"
exon 350..685
BASE COUNT 369 a 218 c 286 g 364 t
ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 1237;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CAAGCCGAGAGATGATG 20
Db 390 CAAGCCGAGAGATGATG 409
RESULT 9
LOCUS AX066402 1374 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 2 from Patent WO0100823.
ACCESSION AX066402
VERSION AX066402.1 GI:12544112
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1374)
AUTHORS Gannon, F., Denger, S. and Flouriot, G.
TITLE Novel isoforms of the human estrogen receptor--g(a)
JOURNAL Patent: WO 0100823-A 2 04-JAN-2001;
EUROPEAN MOLECULAR BIOLOGY LABORATORY (DE)
FEATURES
Source
1. 1374
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 349 a 346 c 386 g 293 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CAAGCCGAGAGATGATG 20
Db 387 CAAGCCGAGAGATGATG 406
RESULT 10
LOCUS S5ESTREC 1788 bp mRNA linear MAM 15-SEP-1994
DEFINITION S.scrofa mRNA for estradiol receptor.
ACCESSION 237167

VERSION 237167.1 GI:587554
KEYWORDS estradiol receptor.
SOURCE Sus scrofa.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 1788)
AUTHORS Bokenkamp, D., Jungblut, P.W. and Thole, H.H.
TITLE The C-terminal half of the porcine estradiol receptor contains no
post-translational modification: determination of the primary
structure
JOURNAL Mol. Cell. Endocrinol. 104 (2), 163-172 (1994)
MEDLINE 95080454
PUBMED 7988744
REFERENCE 2 (bases 1 to 1788)
AUTHORS Bokenkamp, D.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-1994) Dirk Bokenkamp, Max-Planck-Institut fuer
experimentelle Endokrinologie, Feodor-Lynen Strasse 7, Hannover,
30625, Germany
FEATURES
Source
1. 1788
/organism="Sus scrofa"
/db_xref="taxon:9823"
/sex="Female"
/tissue="uterus"
1. 1788
/codon_start=1
/product="estradiol receptor"
/protein_id="CA85524.1"
/db_xref="GI:587555"
/db_xref="SWISS-PROT:O29040"
/translation="MFMTHKTKAGSMALIQIQLANLEPLNPKIPIELPEGVY
DSKRPATYNTPEGAIDFMAAASAVYQSGSLAYGPSESAATGANGLGKQPLNS
VSPSPVLPLAPPQSLPLPHPGQVQVYVLENEPSGVAVREAGPAFRPSNDRGG
GRERLASTSDKSGNAMESEKERTCAVCNDYASGYAVREAGPAFRPSNDRGG
YMCPTAUNOCTIDKNRRKSGQACRLKCYGVAMKGGIRDRGRGMLKHKRDRDGG
RNEVPPGDRSANLPPSPILIKHKNSVYLSLTDQMSILAEAPPIYSYDPT
RPSLEASMGILFTNLADREIVHIMNKKRYPGLDLSHDYHILECAPELIMTIGY
WRSEHFGKILRAPNILLDRNGCYGSEVETIDMLIASSEFRANIOGEEVCLAS
ITLNSGVYTFSLSTKSLERDHIHVLKIDITDILHMAAGATLQOQRRIAQL
LILSHFHMNSKGMELYNMCKNVPLDILIMLDAHRLHAPNIGGPPEDMSOS
QLATSGSTPSHSLOWYITIGEAENFPYTI"
BASE COUNT 410 a 530 c 504 g 344 t
ORIGIN
Query Match 100.0%; Score 20; DB 4; Length 1788;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CAAGCCGAGAGATGATG 20
Db 801 CAAGCCGAGAGATGATG 820
RESULT 11
LOCUS AX411751 1788 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1 from Patent WO0228175.
ACCESSION AX411751
VERSION AX411751.1 GI:21444272
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Chambon, P. and Metzger, D.
TITLE Transgenic mouse for targeted recombination mediated by modified
Cre-er
JOURNAL Patent: WO 0228175-A 1 11-APR-2002;
ASS POOR LE DEV DE LA RECH (FR)
FEATURES
Location/Qualifiers

source 1..1788
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1788
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35258.1"
/db_xref="gi:21444273"
/translation="MTMTLTAKSGMALLHIOGNELEPIINPOLKIPERDLEEVYL
DSSKPAYNPEGAAYEPNAAAAANNOYGGTGLPYRGSAAAPSGNIGGFPPLNS
VSPSPMLIHPROLSPETLOPHGOOVPYTLNBSGTYVBAEPAPYPPNPNDRG
GGERLASTNDGSKAMBSAKETRICVCDYASGHTGVSCGCAFKRSIQSHND
YCPATNQCITDKNRKSCQACLRKCYEVGMAGRIKDRGRMLKHKRQDGE
RCEVGSAGDMRANIMSPSPMTKRSKSLSLTADQVYSLIDAEPIILYSEYDP
RPESSAMGILLTADRELHVMINAKVGEVDLTHOVHLICAMLETIMIGV
WRSMPEKLLFAPNILLDRNGCKVEGVTPEPMLATSSRRPMNLGCEEPVCLKS
IILNSGYTFLSTLKSLENDHHRVIDITTLHLHAKAGITLQOORHLMQEL
LILSHIRMSKMGHEHLSKCKANVPVLDLLEMLDHLRLHAPTSRGASVEETDQS
HLTAGTSVSKSLQKTYTGEAEPVAV"

BASE COUNT 418 a 515 c 508 g 347 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1788;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCGCCAGAGATGATG 20
|||||
DB 801 CAAGCGCCAGAGATGATG 820

RESULT 12
OAE5NREC 1966 bp mRNA linear MAM 08-OCT-1996
LOCUS
DEFINITION O.aries mRNA for estrogen receptor.
ACCESSION Z49257
VERSION Z49257.1 GI:1617201
KEYWORDS estrogen receptor.
SOURCE
ORGANISM Ovis aries.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.
1 (bases 1 to 1966)
Madigou,T., Tiffoche,C., Lazennec,G., Pelletier,J. and
Thieulant,M.L.
The sheep estrogen receptor: cloning and regulation of expression
in the Hypothalamo-pituitary axis
Mol. Cell. Endocrinol. 121 (2), 153-163 (1996)
JOURNAL
MEDLINE 8892316
97047396
PUBMED
REFERENCE 2 (bases 1 to 1966)
Madigou,T.
Direct Submision
Submitted (12-MAY-1995) Thierry Madigou, Interactions Cellulaires
et Moleculaires, CNRS URA 256, Rennes, 35042, FRANCE
FEATURES
location/Qualifiers
1..1966
/organism="Ovis aries"
/strain="Romanov"
/db_xref="taxon:9940"
/clone="lambda p112"
/sex="Female"
/tissue_type="anterior pituitary gland"
/clone_lib="cDNA library"
/dev_stage="adult"
1..1966
/product="estrogen receptor"
BASE COUNT 439 a 604 c 549 g 374 t
ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 1966;
Best Local Similarity 100.0%; Pred. No. 6.9;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCGCCAGAGATGATG 20
|||||
DB 953 CAAGCGCCAGAGATGATG 972

RESULT 13
AX474709 2092 bp DNA PAT 12-AUG-2002
LOCUS
DEFINITION Sequence 9 from Patent WO0222882.
ACCESSION AX474709
VERSION AX474709.1 GI:22214043
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Stanton,M., Epstein,D. and Hamaguchi,N.
TITLE Target activated nucleic acid biosensor and methods of using same
JOURNAL Patent: WO 0222882-A 9 21-MAR-2002;
Archemix Corporation (US)
FEATURES
location/Qualifiers
1..2092
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 473 a 605 c 593 g 421 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2092;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCGCCAGAGATGATG 20
|||||
DB 1093 CAAGCGCCAGAGATGATG 1112

RESULT 14
I08538 2092 bp DNA PAT 02-DEC-1994
LOCUS
DEFINITION Sequence 1 from Patent WO 8705049.
ACCESSION I08538
VERSION I08538.1 GI:588753
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2092)
AUTHORS Shine,J.
TITLE EUCARYOTIC EXPRESSION OF STEROID RECEPTOR PROTEINS
JOURNAL Patent: WO 8705049-A 1 27-AUG-1987;
FEATURES
location/Qualifiers
1..2092
/organism="unknown"
BASE COUNT 473 a 605 c 593 g 421 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2092;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCGCCAGAGATGATG 20
|||||
DB 1093 CAAGCGCCAGAGATGATG 1112

RESULT 15
H0869CF 2092 bp mRNA linear PRI 08-NOV-1994
LOCUS
DEFINITION Human estrogen receptor mRNA, complete cds.
ACCESSION M12674

VERSION M12674.1 GI:182192
KEYWORDS estrogen receptor.
SOURCE Human breast cancer cell line MCF-7, cDNA to mRNA, clone OR8.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2092)
AUTHORS Greene,G.L., Gilna,P., Waterfield,M., Baker,A., Hort,Y. and
Shine,U.
TITLE Sequence and expression of human estrogen receptor complementary
DNA
JOURNAL Science 231 (4742), 1150-1154 (1986)
MEDLINE 86122927
PubMed 3753802
FEATURES
SOURCE Location/Qualifiers
1..2092
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="6q24-q27"
/cell_line="MCF-7"
/cell_type="Mammary Carcinoma"
/lab_host="Prokaryotae"
/tissue_lib="lambda gt11"
1..2092
/gene="ESR"
<1..>2092
/gene="ESR"
/product="ER mRNA"
293..2080
/gene="ESR"
/note="estrogen receptor"
/codon_start=1
/protein_id="AA52399.1"
/db_xref="GI:182193"
/db_xref="GDB:G00-119-120"
/translation="MTMTLHTKASGMALHIOQNELEPLNRPQLPIERPLGEVYL
DSKPAYNTPGCAAYEPNAAANAAYGGTGLPYGPGSEAAAGSNGLGGPPLNS
VSPSLMLTPPOLSPFLPQGOOVPEYLENEBGGYVREAGPPATVRRPNSDRROG
GREPLASTNDKSGMAESAKETRYCAVANDYASGHHGWSCGCKAFKRSIOGND
TMCPTNQCITIDKRRKSCQACRLKRCYEGMMKGGIRKDRSGRRLKHKRRDDGSG
RGEVSGADRRANLNPSPMLTKRSKNSIALSLADQVSAALDAEPILYSEYDPT
RPFESASGMGILTNLADRELVTMINAKRYPGFVDLTHDQVHLFCAMLEILMIGLY
WRSMHPVKLLFAPNLLDRNQKCEGMEVIEPMLLATSRRPMNLQGEFFYGLKS
IILNSGYVTFILSTLKSLEEKDTHRYLDKIDTDLIHMAKAGILIQOHRLAOLL
LISHIRMSKMGKMEIYMKCKNYPIYDILLEMLDAHRLHAPTSRGSASVEETDGS
HIAIRGSTSSHSIQKTYITGKNGPRTV"

BASE COUNT 473 a 605 c 593 g 421 t
ORIGIN 1 bp upstream of EcoRI site.

Query Match 100.0%; Score 20; DB 9; Length 2092;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGATGATG 20
|||||
DB 1093 CAAGCGCCAGAGATG 1112

Search completed: May 17, 2003, 00:56:23
Job time : 406.369 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 12:48:16 ; Search time 111.126 Seconds
(without alignments)
405.306 Million cell updates/sec

Title: US-10-052-092-15

Sequence: . 1 caagcgccagagatgatg 20

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

```
Searched:      2185239 segs, 1125999159 residues
Total number of hits satisfying chosen parameters:  4370478
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Genseq_101002:*

1: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
5: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| c 1 | 20 | 100.0 | 26 | 18 | AAV00167 | Human estrogen re |
| 2 | 20 | 100.0 | 521 | 20 | AAK37808 | PM12 estrogen re |
| 3 | 20 | 100.0 | 539 | 20 | AAK37813 | PM12 estrogen re |
| 4 | 20 | 100.0 | 566 | 20 | AAK37812 | PM14 estrogen re |
| 5 | 20 | 100.0 | 1223 | 22 | AAF29924 | Human estrogen rec |
| 6 | 20 | 100.0 | 1374 | 22 | AAF29925 | Human estrogen rec |
| 7 | 20 | 100.0 | 1380 | 22 | AAC65919 | Nucleotide sequenc |
| 8 | 20 | 100.0 | 1788 | 24 | AB157497 | Human nuclear oest |
| 9 | 20 | 100.0 | 2092 | 8 | AAAT0680 | cDNA encoding huma |

| | | | | | | | |
|----|------|-------|--------|----|-----------|------------------------------------|--------------------|
| 10 | 20 | 100.0 | 2092 | 22 | AAC86930 | Nucleotide sequenc | |
| 11 | 20 | 100.0 | 2092 | 22 | ABL51900 | Human oestrogen re | |
| 12 | 20 | 100.0 | 2220 | 24 | ABA01104 | Fused ER protein c | |
| 13 | 20 | 100.0 | 2322 | 20 | AAK60628 | SSR-LBD fusion pol | |
| 14 | 20 | 100.0 | 4963 | 16 | AAPO6873 | PhCMV α -1 promoter | |
| 15 | 20 | 100.0 | 4963 | 16 | AAK67270 | PhCMV α -1-controlle | |
| 16 | 20 | 100.0 | 4963 | 20 | AAK81725 | cDNA encoding rabb | |
| 17 | 20 | 100.0 | 4963 | 20 | AAK21906 | Rabbit progesteron | |
| 18 | 20 | 100.0 | 4963 | 20 | AAK01370 | PhCMV α -1 rabbit pr | |
| 19 | 20 | 100.0 | 4963 | 22 | AAD09838 | Rabbit progesteron | |
| 20 | 20 | 100.0 | 6450 | 20 | AA223433 | Human estrogen rec | |
| 21 | 20 | 100.0 | 6639 | 22 | AAD06055 | Plasmd α E2CLMBDS | |
| 22 | 20 | 100.0 | 6695 | 22 | AAD06051 | Plasmd α C7LMBDS | |
| 23 | 20 | 100.0 | 6695 | 22 | AAD06057 | Plasmd α C7LMBDSg40 | |
| 24 | 20 | 100.0 | 6695 | 22 | AAD06058 | Plasmd α C7LMBDSG52 | |
| 25 | 20 | 100.0 | 6801 | 22 | AAD06059 | Plasmd α C7LMBDSVP1 | |
| 26 | 20 | 100.0 | 6818 | 22 | AAD06050 | Plasmd α C7LMBd α en | |
| 27 | 20 | 100.0 | 6833 | 22 | AAD06053 | Plasmd α C7LMBDS en | |
| 28 | 20 | 100.0 | 6900 | 22 | AAD06045 | Plasmd α 2C7LMBDS | |
| 29 | 20 | 100.0 | 6956 | 22 | AAD06052 | Plasmd α C7LMBCT en | |
| 30 | 20 | 100.0 | 7038 | 22 | AAD06046 | Plasmd α 2C7LMBDS e | |
| 31 | 20 | 100.0 | 11365 | 24 | AAK87505 | Directed evolution | |
| 32 | 20 | 100.0 | 465237 | 24 | AAQ87681 | Human oestrogen re | |
| 33 | 20 | 100.0 | 465237 | 24 | AAK90193 | Human oestrogen re | |
| 34 | 20 | 100.0 | 465237 | 20 | 16 | AAO85945 | Primer ER-4 to amp |
| 35 | 18 | 90.0 | 550 | 20 | 16 | AAO85945 | Human ORFX ORF403 |
| 36 | 17.4 | 87.0 | 288 | 21 | AAA87501 | Rat hepatocyte car | |
| 37 | 17.4 | 87.0 | 431 | 21 | AAAF1625 | Aspergillus niger | |
| 38 | 16.4 | 82.0 | 564 | 22 | AAH15192 | Human breast cancer | |
| 39 | 16.4 | 82.0 | 568 | 22 | AAI22062 | Human breast cancer | |
| 40 | 16.4 | 82.0 | 1047 | 24 | AAK87184 | Endocrine disrupti | |
| 41 | 16.4 | 82.0 | 1470 | 22 | AAAS06085 | Angiotensin conver | |
| 42 | 16.4 | 82.0 | 1630 | 22 | AAAS06058 | Angiotensin conver | |
| 43 | 16.4 | 82.0 | 2223 | 23 | ABT26609 | Drosophila melanog | |
| 44 | 16.4 | 82.0 | 5088 | 23 | ABT26608 | Drosophila melanog | |
| 45 | 16.4 | 82.0 | 13637 | 23 | ABT16716 | Drosophila melanog | |

ALIGNMENTS

| XX | RESTLT 1 |
|----|--|
| XX | AAV00167/c |
| XX | AAV00167 standard; DNA; 26 BP. |
| XX | AAV00167; |
| XX | 23-MAR-1998 (first entry) |
| XX | Human oestrogen receptor PCR antisense primer E-7. |
| XX | PCR primer: human; progesterone receptor; oestrogen; CD8; glucagon; |
| XX | interleukin 2; insulin; parathyroid hormone; cholecystokinin; KW |
| XX | enkephalin; adrenocorticotrophic hormone; thyroid stimulatory hormone; |
| XX | detection; diagnosis; ss. |
| XX | Synthetic. |
| XX | OS Homo sapiens. |
| XX | JP09187299-A. |
| XX | 22-JUL-1997. |
| XX | 05-JAN-1996; 96JP-0027222. |
| XX | 05-JAN-1996; 96JP-0027222. |
| XX | (NIBI-) NIPON BIOTHERAPY KK. |
| XX | WPI; 1997-419410/39. |
| XX | Polymerase chain reaction primers for detection of mRNA encoding |
| XX | various human protein(s) - e.g. CD8, interleukin-2, parathyroid |

```

PT hormone, adrenocorticotrophic hormone and thyroid stimulating
PT hormone
XX
XX Claim 1; Page 2; 4pp; Japanese.
PS
XX
CC The present sequence represents a novel polymerase chain reaction
CC (PCR) primer. The PCR primers of the present specification are used
CC in the detection of mRNA encoding various human proteins. The mRNA
CC can be detected in a single cell.
XX
SQ Sequence 26 BP; 2 A; 8 C; 5 G; 11 T; 0 other;
Query Match 100.0%; Score 20; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
    |||
    20 CAAGCGCCAGAGAGATGATG 1
DB
RESULT 2
AA37808
ID AAX37808 standard; DNA; 521 BP.
XX
XX AAX37808;
AC
XX
XX 09-JUL-1999 (first entry)
DE PDM12 oestrogen receptor DNA binding domain fusion protein DNA.
XX
XX Peptide display carrier package; PDCP; chimeric protein; fusion protein;
XX screening; library; secreted peptide; oestrogen receptor;
XX DNA binding domain; ss.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 41..475
    /tag= a
    /product= "Oestrogen receptor DNA binding domain
    fusion protein 5'-end"
XX
XX W09911785-A1.
XX
XX 11-MAR-1999.
XX
XX 02-SEP-1998; 98WO-GB02630.
XX
XX 02-SEP-1997; 97GB-0018455.
XX
XX (ROME-) ROWETT RES SERVICES LTD.
XX
XX McGregor D;
XX
XX WPI; 1999-302254/25.
XX
XX P-PSDB; AAY08223.
XX
XX
PT New peptide display carrier package useful for screening nucleotide
PT libraries for sequences that encode peptides of interest
XX
XX Example 1; Fig 1; 103pp; English.
XX
XX This invention describes a novel peptide display carrier package (PDCP),
XX comprising a recombinant polynucleotide-chimeric protein complex, for
XX the isolation of peptide-encoding nucleic acid sequences from a DNA
XX expression library. The PDCP comprises a recombinant
XX polynucleotide-chimeric protein complex, where (a) the chimeric protein
XX has a nucleotide binding portion (NBP) and a target peptide portion
XX (TPP); (b) the recombinant polynucleotide comprises a specific sequence
XX motif which is specifically bound by the NBP; and (c) at least the
XX chimeric protein-encoding portion of the polynucleotide not bound by

```

```

CC the NBP is protected by a binding moiety. The PDCP is used in the
CC method for screening nucleotide libraries for sequences that encode
CC peptides of interest. Unlike prior art, the invention provides a method
CC to produce a peptide library where the DNA is protected from degradation,
CC and which allows production of secreted peptides with a free carboxy
CC terminus. This sequence encodes a fusion protein fragment used in the
CC method of the invention.
XX
SQ Sequence 521 BP; 153 A; 109 C; 155 G; 104 T; 0 other;
Query Match 100.0%; Score 20; DB 20; Length 521;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
    |||
    427 CAAGCGCCAGAGAGATGATG 446
DB
RESULT 3
AA37813
ID AAX37813 standard; DNA; 539 BP.
XX
XX AAX37813;
AC
XX
XX 09-JUL-1999 (first entry)
DE PDM16 oestrogen receptor DNA binding domain fusion protein DNA.
XX
XX Peptide display carrier package; PDCP; chimeric protein; fusion protein;
XX screening; library; secreted peptide; oestrogen receptor;
XX DNA binding domain; ss.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 41..481
    /tag= a
    /product= "Oestrogen receptor fusion protein"
XX
XX W09911785-A1.
XX
XX 11-MAR-1999.
XX
XX 02-SEP-1998; 98WO-GB02630.
XX
XX 02-SEP-1997; 97GB-0018455.
XX
XX (ROME-) ROWETT RES SERVICES LTD.
XX
XX McGregor D;
XX
XX WPI; 1999-302254/25.
XX
XX P-PSDB; AAY08228.
XX
XX
PT New peptide display carrier package useful for screening nucleotide
PT libraries for sequences that encode peptides of interest
XX
XX Example 5; Fig 5; 103pp; English.
XX
XX This invention describes a novel peptide display carrier package (PDCP),
XX comprising a recombinant polynucleotide-chimeric protein complex, for
XX the isolation of peptide-encoding nucleic acid sequences from a DNA
XX expression library. The PDCP comprises a recombinant
XX polynucleotide-chimeric protein complex, where (a) the chimeric protein
XX has a nucleotide binding portion (NBP) and a target peptide portion
XX (TPP); (b) the recombinant polynucleotide comprises a specific sequence
XX motif which is specifically bound by the NBP; and (c) at least the
XX chimeric protein-encoding portion of the polynucleotide not bound by
XX the NBP is protected by a binding moiety. The PDCP is used in the
XX method for screening nucleotide libraries for sequences that encode
XX peptides of interest. Unlike prior art, the invention provides a method

```


CC to produce a peptide library where the DNA is protected from degradation,
CC and which allows production of secreted peptides with a free carboxy
CC terminus. This sequence encodes a fusion protein fragment used in the
CC method of the invention.
XX
XX
SQ Sequence 539 BP; 148 A; 118 C; 165 G; 108 T; 0 other;
Query Match 100.0%; Score 20; DB 20; Length 539;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
DB 385 CAAGCGCCAGAGAGATGATG 404
RESULT 4
AA37812
ID AA37812 standard; DNA; 566 BP.
XX
AC AA37812;
XX
DT 09-JUL-1999 (first entry)
XX
DE PDM14 oestrogen receptor DNA binding domain fusion protein DNA.
XX
KM Peptide display carrier package; PDCP; chimeric protein; fusion protein;
KM screening; library; secreted peptide; oestrogen receptor;
KM DNA binding domain; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 41..475
FT /tag= a
FT /product= "oestrogen receptor fusion protein"
XX
PN MO9911785-A1.
XX
PD 11-MAR-1999.
XX
PF 02-SEP-1998; 98WO-GB02630.
XX
PR 02-SEP-1997; 97GB-0018455.
XX
PA (ROME-) ROWETT RES SERVICES LTD.
XX
PI McGregor D;
XX
DR WPI; 1999-302254/25.
DR P-PSDB; AAY08227.
XX
PT New peptide display carrier package useful for screening nucleotide
PT libraries for sequences that encode peptides of interest
XX
PS Example 4; Fig 4; 103pp; English.
XX
CC This invention describes a novel peptide display carrier package (PDCP),
CC comprising a recombinant polynucleotide-chimeric protein complex, for
CC the isolation of peptide-encoding nucleic acid sequences from a DNA
CC expression library. The PDCP comprises a recombinant
CC polynucleotide-chimeric protein complex, where (a) the chimeric protein
CC has a nucleotide binding portion (NBP) and a target peptide portion
CC (TPP); (b) the recombinant polynucleotide comprises a specific sequence
CC motif which is specifically bound by the NBP; and (c) at least the
CC chimeric protein-encoding portion of the polynucleotide not bound by
CC the NBP is protected by a binding moiety. The PDCP is used in the
CC method for screening nucleotide libraries for sequences that encode
CC peptides of interest. Unlike prior art, the invention provides a method
CC to produce a peptide library where the DNA is protected from degradation,
CC and which allows production of secreted peptides with a free carboxy
CC terminus. This sequence encodes a fusion protein fragment used in the

CC method of the invention.
XX
XX
SQ Sequence 566 BP; 164 A; 118 C; 169 G; 115 T; 0 other;
Query Match 100.0%; Score 20; DB 20; Length 566;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
DB 427 CAAGCGCCAGAGAGATGATG 446
RESULT 5
AAF29924
ID AAF29924 standard; DNA; 1223 BP.
XX
AC AAF29924;
XX
DT 04-APR-2001 (first entry)
XX
DE Human estrogen receptor alpha isoform #1 DNA.
XX
KM Human; estrogen receptor alpha; cancer; osteoporosis; bone;
KM Alzheimer's; cardiovascular; ds.
XX
OS Homo sapiens.
XX
PN WO200100823-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-EP05981.
XX
PR 29-JUN-1999; 99IT-WI01433.
XX
PA (EURO-) EURO MOLECULAR BIOLOGY LAB.
XX
PI Gannon F, Dengler S, Flouriot G;
XX
DR WPI; 2001-137955/14.
XX
XX
PT Novel isoforms of human estrogen receptor alpha useful for preparing
PT therapeutic agents for treating cancer, osteoporosis, Alzheimer's
PT disease and cardiovascular diseases
XX
PS Claim 1; Pag e44; 53pp; English.
XX
CC The present invention relates to a human estrogen receptor
CC (HER)-alpha isoform. Molecules which modulate the activity of
CC the estrogen receptor are useful for the preparation of therapeutic
CC agents for treating cancer, osteoporosis and other bone disorders,
CC Alzheimer's disease and cardiovascular diseases.
XX
SQ Sequence 1223 BP; 303 A; 316 C; 341 G; 263 T; 0 other;
Query Match 100.0%; Score 20; DB 22; Length 1223;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
DB 236 CAAGCGCCAGAGAGATGATG 255
RESULT 6
AAF29925
ID AAF29925 standard; DNA; 1374 BP.
XX
AC AAF29925;
XX
DT 04-APR-2001 (first entry)
XX

DE Human estrogen receptor alpha isoform #2 DNA.
XX
KW Human; estrogen receptor alpha; cancer; osteoporosis; bone;
XX Alzheimer's; cardiovascular; ds.
OS Homo sapiens.
XX
PN WO200106823-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-EP05981.
XX
PR 29-JUN-1999; 99IT-MT01433.
XX
PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX
PI Gannon F, Dengler S, Flouriot G;
XX
DR WPI, 2001-137955/14.
XX
PT Novel isoforms of human estrogen receptor alpha useful for preparing
PT therapeutic agents for treating cancer, osteoporosis, Alzheimer's
PT disease and cardiovascular diseases -
XX
PS Claim 1; Page 44-45; 53pp; English.
XX
CC The present invention relates to a human estrogen receptor
CC (HER)-alpha isoform. Molecules which modulate the activity of
CC the estrogen receptor are useful for the preparation of therapeutic
CC agents for treating cancer, osteoporosis and other bone disorders,
CC Alzheimer's disease and cardiovascular diseases.
XX
SQ Sequence 1374 BP; 349 A; 346 C; 386 G; 293 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 1374;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGGCCAGAGATGATG 20
DB 387 CAAGGCCAGAGATGATG 406
|||||

RESULT 7
AAC86919
ID AAC86919 standard; DNA; 1380 BP.
XX
AC AAC86919;
XX
DT 02-APR-2001 (first entry)
XX
DE Nucleotide sequence of the DNA binding domain of estrogen receptor.
XX
KW Ribozyme; oestrogen-dependent tumour; cell proliferation; glucocorticoid;
KW DNA-binding domain; oestrogen receptor; cancer treatment; breast cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200074485-A1.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US15243.
XX
PR 04-JUN-1999; 99US-0137470.
XX
PA (TEXA) UNIV TEXAS.
XX
PI Roy AK, Lavyrovsky Y, Tyagi RK, Song CS, Chatterjee B;
XX WPI, 2001-061633/07.
XX

XX
PT Ribozyme having a high substrate specificity for an mRNA encoding a
PT DNA-binding domain of human estrogen receptor, useful for inhibiting
PT estrogen-dependent tumour cell proliferation, particularly breast cancer
PT
XX
PS Claim 1; Page 6-7; 49pp; English.
XX
CC The specification describes a ribozyme capable of inhibiting
CC oestrogen-dependent tumour cell proliferation and having a high
CC substrate specificity for an mRNA sequence encoding a DNA-binding
CC domain of human estrogen receptor. The ribozyme is free of endonuclease
CC activity for an mRNA having a DNA binding domain of a glucocorticoid.
CC The oestrogen receptor site-specific ribozymes are useful for cancer
CC treatment and therapies, especially for inhibiting oestrogen-dependent
CC tumour cell proliferation, particularly breast cancer. The present
CC sequence represents the DNA binding domain of a human oestrogen receptor.
XX
SQ Sequence 1380 BP; 320 A; 385 C; 400 G; 275 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 1380;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGGCCAGAGATGATG 20
DB 733 CAAGGCCAGAGATGATG 752
|||||

RESULT 8
ABL57497
ID ABL57497 standard; cDNA; 1788 BP.
XX
AC ABL57497;
XX
DT 22-NOV-2002 (first entry)
XX
DE Human nuclear oestrogen receptor alpha coding sequence.
XX
KW Oestrogen; receptor; human; transgenic mouse; cytosolic;
KW antiinflammatory; antidiabetic; endocrine; anorectic; hepatotropic;
KW gene; ss.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FH 1..1788
FH CDS /tag= a
FH /product= "Nuclear oestrogen receptor"
FT
FT
XX
PN WO200228175-A2.
XX
PD 11-APR-2002.
XX
PF 28-SEP-2001; 2001WO-IB02246.
XX
PR 03-OCT-2000; 2000FR-0012570.
XX
PR 11-MAY-2001; 2001US-0853033.
XX
PA (ADER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.
XX
PI Chambon P, Metzger D;
XX
DR WPI, 2002-463217/49.
XX
DR P-PSDB; ABB76378.
XX
PT A transgenic mouse, useful in screening for medicaments for the
PT treatment of e.g., diabetes or skin cancers, comprises a fusion protein
PT between a recombinase Cre, and a modified ligand binding domain of the
PT nuclear estrogen receptor alpha
XX
PS Disclosure; Page 128-130; 149pp; English.
XX

CC The present sequence is the coding sequence of the human nuclear
CC oestrogen receptor alpha (1). The invention relates to a non-human
CC metazoan organism, especially a transgenic mouse, characterised in
CC that at least one cell comprises: (i) a fusion protein formed from
CC a recombinase (Cre), a hinge region (preferably derived from human
CC (1)), and a modified ligand binding domain (LBD) of a nuclear
CC oestrogen receptor, such as (1), its fragment or variant; and (ii)
CC one or more genes or DNA sequences of interest belonging to the
CC genome of the organism, into which one or more recognition sites
CC of the recombinase protein are inserted. The (i) LBD domain is
CC preferably modified by a G521R, G400V, or M543A/V544A mutation.
CC The fusion protein has negligible, or even zero, recombinase
CC activity in the presence of a natural ligand such as oestradiol;
CC but recombinase activity is induced by a small quantity of a
CC synthetic ligand that has antioestrogenic activity, e.g. tamoxifen
CC or 4-hydroxytamoxifen. The metazoan organism or its cells, such
CC as epidermal cells, hepatocytes or adipocytes, are useful in
CC carrying out a spatiotemporally controlled site-specific
CC recombination of a DNA sequence of interest in its natural
CC chromatin environment. It is also used in screening of medicaments
CC for pathological conditions associated with an alteration of the
CC expression and/or function of the DNA sequence of interest, such as
CC skin cancer, inflammation, diabetes, alopecia, obesity, or in
CC promoting hepatic regeneration.

SO Sequence 1788 BP; 418 A; 515 C; 508 G; 347 T; 0 other;
Query Match 100.0%; Score 20; DB 24; Length 1788;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGGCCAGAGATGATG 20
|||||
Db 801 CAAGGCCAGAGATGATG 820

RESULT 9
ID AANT0880 standard; cDNA; 2092 BP.

XX AANT0880;
XX 24-FEB-1991 (first entry)
XX cDNA encoding human oestrogen receptor protein.
XX DE
XX KM Human oestrogen receptor protein; ss cDNA; expression system; assay;
XX KW steroid receptor protein;
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 293..1280
XX FT /*Tag= a
XX PN W08705049-A.
XX PD 27-AUG-1987.
XX PF 18-FEB-1987; 87WO-US00341.
XX PR 20-FEB-1986; 86US-0833829.
XX PA (CALB-) CALIFORNIA BIOTECHNOLOGY INC.
XX PI Shine J;
XX WPI; 1987-250213/35.
XX DR P-PSDB; AAP70543.
XX Expression system for vertebrate steroid receptor protein - comprising
XX PT DNA sequence encoding the protein linked to control sequences in
XX eucaryotic hosts.

XX Disclosure; Fig. 1-1 - 1-2; 26pp; English.

XX The cDNA can be used in an expression system to express human
XX oestrogen receptor protein. The coding sequence is operably linked to
XX control sequences compatible with eukaryotic host cells. This method
XX allows expression under conditions which favour appropriate post-
XX translational processing. It produces large amts. of purified protein
XX useful in the design of agonist and antagonist cpts., for the study of
XX the mechanism of action of the steroid binding proteins in general,
XX and for use in diagnostic assays for the proteins or antibodies to
XX them. These assays are important in, eg the diagnosis of tumour
XX sensitivities to steroid metabolism. Suitable host cells are VERO,
XX HeLa and CHO cells.

SO Sequence 2092 BP; 473 A; 607 C; 591 G; 421 T; 0 other;
Query Match 100.0%; Score 20; DB 8; Length 2092;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGGCCAGAGATGATG 20
|||||
Db 1093 CAAGGCCAGAGATGATG 1112

RESULT 10
ID AAC86920 standard; cDNA; 2092 BP.

XX AAC86920;
XX 02-APR-2001 (first entry)
XX Nucleotide sequence of the human oestrogen receptor cDNA.
XX DE
XX KM Ribozyme; oestrogen-dependent tumour; cell proliferation; glucocorticoid;
XX KM DNA-binding domain; oestrogen receptor; cancer treatment; breast cancer;
XX KW ss.
XX OS Homo sapiens.
XX PN W0200074485-A1.
XX PD 14-DEC-2000.
XX PF 02-JUN-2000; 2000WO-US15243.
XX PR 04-JUN-1999; 99US-0137470.
XX PA (TEXA) UNIV TEXAS.
XX PI Roy AK, Lavyrovsky Y, Tyagi RK, Song CS, Chatterjee B;
XX WPI; 2001-061633/07.
XX Ribozyme having a high substrate specificity for an mRNA encoding a
XX PT DNA-binding domain of human estrogen receptor, useful for inhibiting
XX PT estrogen-dependent tumour cell proliferation, particularly breast cancer
XX -
XX Disclosure; Page 8-9; 49pp; English.
XX The specification describes a ribozyme capable of inhibiting
XX oestrogen-dependent tumour cell proliferation and having a high
XX substrate specificity for an mRNA sequence encoding a DNA-binding
XX domain of human oestrogen receptor. The ribozyme is free of endonuclease
XX activity for an mRNA having a DNA binding domain of a glucocorticoid.
XX The oestrogen receptor site-specific ribozymes are useful for cancer
XX treatment and therapies, especially for inhibiting oestrogen-dependent
XX tumour cell proliferation, particularly breast cancer. The present
XX sequence represents the human oestrogen receptor cDNA.

SO Sequence 2092 BP; 473 A; 605 C; 593 G; 421 T; 0 other;
Query Match 100.0%; Score 20; DB 24; Length 2092;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
DB 1093 CAAGCGCCAGAGAGATGATG 1112
RESULT 11
ID ABL51900 standard; cDNA: 2092 BP.
AC ABL51900;
XX ABL51900;
XX ABL51900;
XX 10-JUL-2002 (first entry)
XX Human oestrogen receptor encoding cDNA SEQ ID No.9.
XX
XX Target activated nucleic acid biosensor; signalling moiety; human;
XX nucleic acid sensor; detection; engineering; drug optimisation;
XX oestrogen receptor; gene; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 293..2080
XX /tag= a
XX /product= "oestrogen receptor"
XX
XX WO200222882-A2.
XX
XX 21-MAR-2002.
XX
XX 13-SEP-2001; 2001MO-US28835.
XX
XX 13-SEP-2000; 2000US-232454P.
XX
XX (ARCH-) ARCHEMIX CORP.
XX
XX Stanton M, Epstein D, Hamaguchi N;
XX WPI: 2002-393977/42.
XX P-PSDB; ABB09265.
XX
XX Nucleic acid sensor for detecting target molecule, comprises target
XX molecule activation site and optical signalling unit that changes its
XX optical properties upon allosteric modulation sensor after recognition
XX of target -
XX
XX Example 2; Page 75-76; 144pp; English.
XX
XX The present invention describes a nucleic acid sensor molecule (I)
XX comprising a target molecule activation site comprising a structure that
XX recognises a target molecule and an optical signalling unit including at
XX least one nucleotide coupled to a signalling moiety that changes its
XX optical properties upon allosteric modulation of (I) following
XX recognition of the target molecule. (I) is useful for detecting a target
XX molecule associated with a pathological condition or genetic alteration.
XX (I) is useful for identifying a drug compound, by identifying a nucleic
XX acid biosensor-based molecular profile of target molecules associated with
XX a disease trait in a patient, administering a candidate compound to the
XX patient, and monitoring changes in the profile. Alternatively, the method
XX involves identifying a number of pathway target molecules, administering
XX a candidate compound to a patient having a disease trait, and monitoring
XX changes in the structure, level or activity of two or more of the pathway
XX target molecules using (I). The profile of target molecules or the
XX changes in the structure is compared to the profile of a reference
XX healthy or diseased population. (I) is useful in multiple assays, for the
XX detection of target molecule. (I) is also useful in diagnostic
XX applications and drug optimisation. The present sequence encodes the

CC human oestrogen receptor, which is used in an example from the present
XX invention.
XX
XX Sequence 2092 BP; 473 A; 605 C; 593 G; 421 T; 0 other;
SO
Query Match 100.0%; Score 20; DB 24; Length 2092;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
DB 1093 CAAGCGCCAGAGAGATGATG 1112
RESULT 12
ID ABA01104 standard; cDNA: 2220 BP.
AC ABA01104;
XX ABA01104;
XX ABA01104;
XX 01-FEB-2002 (first entry)
XX Fused ER protein cDNA.
XX
XX Fused ER protein cDNA.
XX
XX Human; oestrogen receptor; ER; ERalpha; fusion; screening;
XX sugar-binding protein; ss.
XX
XX Chimeric - Homo sapiens.
XX Chimeric - Unidentified.
XX JP2001258572-A.
XX
XX 25-SEP-2001.
XX
XX 23-MAR-2000; 2000JP-0082034.
XX
XX 23-MAR-2000; 2000JP-0082034.
XX
XX (TOYMA) TOYOMA KK.
XX
XX WPI: 2002-003334/01.
XX P-PSDB; AAM51591.
XX
XX Fusion protein for screening exogenous increase disturbance
XX substances, comprises the fused estrogen receptor protein -
XX
XX Claim 10; Page 9-10; 14pp; Japanese.
XX
XX The invention relates to a fused oestrogen receptor (ER) protein in
XX which an ER protein is fused with a sugar-binding protein through a
XX protease-resistant sequence. The fused ER protein can be used for
XX screening exogenous increase disturbance substances. A
XX polynucleotide encoding ERalpha was isolated from human cDNA by PCR.
XX The present sequence encodes the fusion protein of the invention.
XX
XX Sequence 2220 BP; 594 A; 562 C; 604 G; 460 T; 0 other;
SO
Query Match 100.0%; Score 20; DB 24; Length 2220;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGAGATGATG 20
DB 1236 CAAGCGCCAGAGAGATGATG 1255
RESULT 13
ID AAX60628 standard; DNA: 2322 BP.
AC AAX60628;
XX AAX60628;
XX 03-AUG-1999 (first entry)

```

XX DE SSR-LBD fusion polypeptide encoding DNA.
XX XX
XX KW Fusion protein; recombinase; nuclear receptor; ligand binding domain;
XX KW site-specific recombination; SSR-LBD fusion protein; FLP recombinase;
XX KW estrogen binding domain; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT CDS 1..2322
XX FT /tag= a
XX FT /product= "SSR-LBD fusion protein"
XX FT misc_feature 1..1269
XX FT /tag= b
XX FT /note= "sequence coding for FLP recombinase domain"
XX FT misc_feature 1270..1284
XX FT /tag= c
XX FT /note= "sequence coding for the linker peptide"
XX FT misc_feature 1285..2322
XX FT /tag= d
XX FT /note= "sequence coding for estrogen binding domain"
XX PN WO9500555-A1.
XX PD 05-JAN-1995.
XX PF 28-JUN-1994; 94MO-EP02088.
XX PR 28-JUN-1993; 93EP-0110298.
XX PA (EMBL-) EMBL EURO MOLECULAR BIOLOGY LAB.
XX PI Stewart F;
XX DR WPI, 1999-337364/28.
XX DR P-Psdb; AAY21544.
XX PT New fusion protein - comprising a recombinase protein and nuclear
XX PT receptor, useful for regulation of recombination
XX PS Disclosure; Fig 4; 28pp; English.
XX CC The invention describes a fusion protein that comprises a recombinase
XX CC protein or a component of a recombinase complex fused to part of or all
XX CC of a nuclear receptor, where recombinase activity is inhibited in the
XX CC absence of amino acids binding the ligand binding domain of the receptor,
XX CC and activity is altered or induced by ligand binding. The new fusion
XX CC protein is useful for regulating site-specific recombination in cells or
XX CC organisms, for determining the effect of mutations in the ligand binding
XX CC domain of a nuclear receptor on its ability to bind a ligand, and for
XX CC assessing binding of a ligand to the ligand binding domain. Prior art
XX CC methods of recombinase regulation rely on inducible heat-shock promoters,
XX CC but these are only applicable in flies, and no suitable counterpart has
XX CC been found. The present sequence represents a nucleotide sequence coding
XX CC for the SSR-LBD fusion polypeptide. The fusion protein consists of an FLP
XX CC recombinase domain linked by a linker peptide to an estrogen binding
XX CC domain.
XX SQ Sequence 2322 BP; 703 A; 505 C; 545 G; 569 T; 0 other;

Query Match 100.0%; Score 20; DB 20; Length 2322;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGGCCAGAGAGATGATG 20
DB 1335 CAAGGCCAGAGAGATGATG 1354

```

RESULT 14
AAT06873
ID AAT06873 standard; DNA; 4963 BP.

```

XX AC AAT06873;
XX DT 13-MAR-1996 (first entry)
XX DE PhCMW*-1 promoter and human oestrogen receptor.
XX KW Transactivator; tTA; tet operator; tetr; gene expression;
XX KW tetracycline-responsive promoter; PhCMW*-1; human cytomegalovirus;
XX KW transcription activation; eukaryotic cell; estrogen receptor;
XX KW ds; ss.
XX OS Chimeric synthetic;
XX OS Chimeric Homo sapiens.
XX PN US5464758-A.
XX PD 07-NOV-1995.
XX PF 14-JUN-1993; 93US-0076726.
XX PR 14-JUN-1993; 93US-0076726.
XX PA (BUTJ/) BUTJARD H.
XX PA (GOSSE/) GOSSEN M.
XX PI Butjard H, Gossen M;
XX DR WPI, 1995-392612/50.
XX PT Polynucleotide encoding transactivator fusion protein contg. tet
XX PT repressor - used to control expression of gene regulated by minimal
XX PT promoter linked to tet operon, and vectors and cells where gene
XX PT expression is regulated by tetracycline
XX PS Disclosure; Fig 10A-B; 37pp; English.
XX CC CDNA coding for the human oestrogen receptor was placed under
XX CC control of the PhCMW*-1 promoter (see AAT06869), resulting in
XX CC tetracycline-regulated expression vector pUHPgr-3 contg. the
XX CC sequence given in AAT06873. The construct was transfected into
XX CC eukaryotic cells expressing the tTA transactivator (AAT06867) to
XX CC allow tetracycline-regulated expression of the receptor.
XX SQ Sequence 4963 BP; 1259 A; 1300 C; 1278 G; 1126 T; 0 other;

Query Match 100.0%; Score 20; DB 16; Length 4963;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGGCCAGAGAGATGATG 20
DB 1266 CAAGGCCAGAGAGATGATG 1285

```

RESULT 15
AAT06270
ID AAT06270 standard; DNA; 4963 BP.
AC AAT06270;
DT 17-JUL-1995 (first entry)
DE PhCMW*-1-controlled progesterone receptor.
KW tTA; transactivator; tetracycline-controllable transactivator;
KW conditional inactivation; homologous recombination; gene expression;
KW gene regulation; gene therapy; tetracycline-resistance; tetr;
KW transgenic animal; PhCMW*-1; promoter; tetr; CMV;
KW progesterone receptor; ds.
OS Human cytomegalovirus Towne.

PN W09429442-A.
 XX
 PD 22-DEC-1994.
 XX
 PF 14-JUN-1994; 94MO-US06734.
 XX
 PR 14-JUN-1993; 93US-0076327.
 XX
 PA (BADI) BASF AG.
 XX
 PI Bujard H, Gossen M, Salfeld JG, Voss JW;
 XX
 DR WPI; 1995-036472/05.
 XX
 XX
 PT Regulatory systems using tetracycline-controllable transactivator
 PT (ttr) - useful for conditional inactivation or modulation of
 PT gene expression in a host cell or animal
 XX
 XX
 PS Disclosure; Page 54-58; 103pp; English.
 XX
 CC The minimal promoter PhCMV*-1 is based on the human CMV
 CC intermediate-early promoter (with enhancer region removed), fused
 CC to a tetracycline operator, tetrO. The promoter is used to
 CC express tTA transactivator in host cells. In clone pUHD BGR4
 CC (given in AAQ76269), cDNA for rabbit progesterone receptor is under
 CC the control of PhCMV*-1.
 XX
 SO Sequence 4963 BP; 1259 A; 1298 C; 1280 G; 1126 T; 0 other;
 Query Match 100.0%; Score 20; DB 16; Length 4963;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAAGGCGCAGAGAGATGATG 20
 Db 1266 CAAGGCGCAGAGAGATGATG 1285

Search completed: May 16, 2003, 15:58:25
 Job time : 116.126 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compygen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2003, 19:53:58 ; Search time 23.1788 Seconds
(without alignments)
264.618 Million cell updates/sec

Title: US-10-052-092-15

Sequence: 20
1 caagcgcagagagatgatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | ID | Description |
|------------|-------|--------------|---------|---------------------------------------|
| 1 | 20 | 100.0 | 2222 3 | US-08-564-264-2 Sequence 2, Appl |
| 2 | 20 | 100.0 | 4963 1 | US-08-076-726-16 Sequence 16, Appl |
| 3 | 20 | 100.0 | 4963 2 | US-08-260-452-9 Sequence 9, Appl |
| 4 | 20 | 100.0 | 4963 2 | US-08-481-970-9 Sequence 9, Appl |
| 5 | 20 | 100.0 | 4963 2 | US-08-897-719-9 Sequence 9, Appl |
| 6 | 20 | 100.0 | 4963 4 | US-09-163-269-9 Sequence 9, Appl |
| 7 | 20 | 100.0 | 6450 4 | US-09-041-886-34 Sequence 34, Appl |
| 8 | 20 | 100.0 | 6450 4 | US-08-453-998-1 Sequence 1, Appl |
| 9 | 15.8 | 79.0 | 2328 4 | US-08-688-649-38 Sequence 38, Appl |
| 10 | 15.8 | 79.0 | 3499 1 | US-07-966-278-2 Sequence 2, Appl |
| 11 | 15.8 | 79.0 | 3499 1 | US-08-424-921-2 Sequence 2, Appl |
| 12 | 15.8 | 79.0 | 3499 2 | US-08-556-355A-2 Sequence 2, Appl |
| 13 | 15.8 | 79.0 | 3499 2 | US-07-803-627A-2 Sequence 2, Appl |
| 14 | 15.8 | 79.0 | 3499 2 | US-09-089-879-9 Sequence 6, Appl |
| 15 | 15.4 | 77.0 | 3640 4 | US-09-221-017B-1047 Sequence 1047, Ap |
| 16 | 15.4 | 77.0 | 35060 3 | US-08-814-095-7 Sequence 7, Appl |
| 17 | 15.4 | 76.0 | 737 2 | US-08-861-549-2 Sequence 2, Appl |
| 18 | 15.2 | 76.0 | 1308 1 | US-08-518-878B-57 Sequence 57, Appl |
| 19 | 15.2 | 76.0 | 1308 2 | US-08-807-861A-57 Sequence 57, Appl |
| 20 | 15.2 | 76.0 | 1308 3 | US-09-210-681-57 Sequence 57, Appl |
| 21 | 15.2 | 76.0 | 1308 3 | US-08-946-719A-57 Sequence 57, Appl |
| 22 | 15.2 | 76.0 | 3870 1 | US-08-138-641-1 Sequence 1, Appl |
| 23 | 15.2 | 76.0 | 3870 1 | US-08-138-133-1 Sequence 1, Appl |
| 24 | 15.2 | 76.0 | 3893 1 | US-08-138-641-3 Sequence 3, Appl |
| 25 | 15.2 | 76.0 | 3893 1 | US-08-138-133-3 Sequence 3, Appl |
| 26 | 15.2 | 76.0 | 4198 4 | US-09-586-719-11 Sequence 11, Appl |
| 27 | 15.2 | 76.0 | | |

| | | | | |
|----|------|------|---------|-------------------------------------|
| 28 | 14.8 | 74.0 | 575 4 | US-09-385-982-450 Sequence 450, App |
| 29 | 14.8 | 74.0 | 1776 4 | US-09-149-476-59 Sequence 59, Appl |
| 30 | 14.8 | 74.0 | 1791 4 | US-09-149-476-226 Sequence 226, App |
| 31 | 14.8 | 74.0 | 5455 1 | US-08-342-930-1 Sequence 1, Appl |
| 32 | 14.8 | 74.0 | 9848 4 | US-09-385-222A-3 Sequence 3, Appl |
| 33 | 14.8 | 74.0 | 10163 2 | US-08-659-251-1 Sequence 1, Appl |
| 34 | 14.8 | 74.0 | 10163 4 | US-09-256-490-1 Sequence 1, Appl |
| 35 | 14.8 | 74.0 | 10163 5 | PCR-US96-11445-1 Sequence 1, Appl |
| 36 | 14.4 | 72.0 | 810 4 | US-09-230-041-1 Sequence 1, Appl |
| 37 | 14.2 | 71.0 | 30 3 | US-09-167-717-6 Sequence 6, Appl |
| 38 | 14.2 | 71.0 | 750 3 | US-09-167-717-4 Sequence 4, Appl |
| 39 | 14.2 | 71.0 | 864 4 | US-09-439-261-12 Sequence 12, Appl |
| 40 | 14.2 | 71.0 | 864 4 | US-09-227-613-13 Sequence 13, Appl |
| 41 | 14.2 | 71.0 | 918 4 | US-09-439-261-5 Sequence 5, Appl |
| 42 | 14.2 | 71.0 | 918 4 | US-09-227-613-5 Sequence 5, Appl |
| 43 | 14.2 | 71.0 | 960 4 | US-09-439-261-36 Sequence 36, Appl |
| 44 | 14.2 | 71.0 | 960 4 | US-09-227-613-35 Sequence 35, Appl |
| 45 | 14.2 | 71.0 | 990 4 | US-09-439-261-35 Sequence 35, Appl |

ALIGNMENTS

RESULT 1
US-08-564-264-2
; Sequence 2, Application US/08564264
; Patent No. 6040430
; GENERAL INFORMATION:
; APPLICANT: STEWART, FRANCIS
; TITLE OF INVENTION: REGULATION OF SITE-SPECIFIC
; TITLE OF INVENTION: RECOMBINATION BY SITE-SPECIFIC
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Nikaide, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,264
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02088
; FILING DATE: 28-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93 110 298.2
; FILING DATE: 28-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P564-5019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2322 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1-1269)
; OTHER INFORMATION: /note="FLP recombinase domain."

```

; NAME/KEY: misc.feature
; LOCATION: (1270-1284)
; OTHER INFORMATION: /note= "Linker peptide."
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1285-2322)
; OTHER INFORMATION: /note= "Estrogen binding domain."
US-08-564-264-2

Query Match
Best Local Similarity 100.0%; Score 20; DB 3; Length 2322;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGATG 20
DB 1335 CAAGCGCCAGAGAGATGATG 1354

RESULT 2
US-08-076-726-16
; Sequence 16; Application US/08076726
; Patent No. 5464758
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Tight Control of Gene Expression in
; TITLE OF INVENTION: Eucaryotic Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: United States of America
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,726
; FILING DATE: 14-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942,2490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-076-726-16

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 4963;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGATG 20
DB 1266 CAAGCGCCAGAGAGATGATG 1285

RESULT 3
US-08-260-452-9
; Sequence 9; Application US/08260452
; Patent No. 5650298

; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; APPLICANT: Salfeld, Jochem
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
; TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahnive & Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/260,452
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076,327
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-013CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Human cytomegalovirus
; IMMEDIATE SOURCE:
; CLONE: pTHD BGR4
US-08-260-452-9

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 4963;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGATG 20
DB 1266 CAAGCGCCAGAGAGATGATG 1285

RESULT 4
US-08-481-970-9
; Sequence 9; Application US/08481970
; Patent No. 5859310
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; APPLICANT: Salfeld, Jochem
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahnive & Cockfield
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
```



```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.970
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/260.452
FILING DATE: 14-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076.327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4963 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
IMMEDIATE SOURCE:
CLONE: pUHD BGR4
US-08-481-970-9

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 4963;
Pred. No. 0.58;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 1266 CAAGCGCCAGAGAGATGATG 1285

RESULT 5
US-08-897-719-9
Sequence 9, Application US/08897719
Patent No. 5922927
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bajard, Hermann
APPLICANT: Salfeld, Jochem
APPLICANT: Voss, Jeffrey
TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,719
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER: 08/260.452
FILING DATE:
APPLICATION NUMBER: 08/076.327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4963 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
IMMEDIATE SOURCE:
CLONE: pUHD BGR4
US-08-897-719-9

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 4963;
Pred. No. 0.58;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 1266 CAAGCGCCAGAGAGATGATG 1285

RESULT 6
US-09-163-269-9
Sequence 9, Application US/09163269
Patent No. 6252136
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bajard, Hermann
APPLICANT: Salfeld, Jochem
APPLICANT: Voss, Jeffrey
TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,269
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/481,970
FILING DATE:
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
Transcr
```

;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4963 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: circular
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Human cytomegalovirus
;; IMMEDIATE SOURCE:
;; CLONE: PCHD B6R4
;; US-09-163-269-9

Query Match 100.0%; Score 20; DB 4; Length 4963;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCCGACAGAGATGATG 20
1266 CAAGCCGACAGAGATGATG 1285

RESULT 7
US-09-041-886-34
; Sequence 34, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LT 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 361..2146
; US-09-041-886-34

Query Match 100.0%; Score 20; DB 4; Length 6450;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCCGACAGAGATGATG 20
|||||

Db 1161 CAAGCCGACAGAGATGATG 1180

RESULT 8
US-08-453-998-1
; Sequence 1, Application US/08453998
; Patent No. 6444438
; GENERAL INFORMATION:
; APPLICANT: CHABON, PIERRE
; APPLICANT: METZGER, DANIEL
; APPLICANT: WHITE, JOHN
; TITLE OF INVENTION: METHOD FOR THE PREPARATION OF A PROTEIN
; TITLE OF INVENTION: BY YEASTS USING AN INDUCIBLE SYSTEM, VECTORS AND
; TITLE OF INVENTION: CORRESPONDING TRANSFORMED STRAINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: CUSHMAN, DAREY & CUSHMAN
; STREET: 1100 NEW YORK AVE., N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,998
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/161,064
; FILING DATE: 03-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CHAPIN, MARIANA K.
; REGISTRATION NUMBER: 35,843
; REFERENCE/DOCKET NUMBER: 1037/98493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3711
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-453-998-1

Query Match 100.0%; Score 20; DB 4; Length 6450;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCCGACAGAGATGATG 20
Db 1161 CAAGCCGACAGAGATGATG 1180

RESULT 9
US-08-688-649-38
; Sequence 38, Application US/08688649
; Patent No. 5827716
; GENERAL INFORMATION:
; APPLICANT: MAMONE, JOSEPH A.
; TITLE OF INVENTION: MODIFIED POL-II TYPE DNA
; TITLE OF INVENTION: POLYMERASES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: LYON & LYON
; STREET: 633 West Fifth Street
; STREET: Suite 4700

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688,649
FILING DATE: Filed Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/281
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 2328 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-688-649-38

Query Match 79.0%; Score 15.8; DB 1; Length 2328;
Best Local Similarity 89.5%; Pred. No. 57;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAAGCGCCAGAGAGATGAT 19
||||| |||||||
DB 554 CAAGCGAGAGAGATGAT 572

RESULT 10
US-09-715-524B-1
Sequence 1, Application US/09715524B
Patent No. 6333183
GENERAL INFORMATION:
APPLICANT: Evans, Steven
APPLICANT: Mamone, Joseph
APPLICANT: Davis, Maria
APPLICANT: Connolly, Bernard
TITLE OF INVENTION: Improving Dideoxynucleotide-Triphosphate Utilization by
TITLE OF INVENTION: The Hyper-Thermophilic DNA Polymerase from the Archaeon
FILE REFERENCE: PB-9965
CURRENT APPLICATION NUMBER: US/09/715,524B
CURRENT FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/167,066
PRIOR FILING DATE: 1999-11-23
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1
LENGTH: 2328
TYPE: DNA
ORGANISM: Pyrococcus furiosus
US-09-715-524B-1

Query Match 79.0%; Score 15.8; DB 4; Length 2328;
Best Local Similarity 89.5%; Pred. No. 57;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAAGCGCCAGAGAGATGAT 19
||||| |||||||

DB 554 CAAGCGAGAGAGATGAT 572

RESULT 11
US-07-966-278-2
Sequence 2, Application US/07966278
Patent No. 5489523
GENERAL INFORMATION:
APPLICANT: Mathur, Eric A
TITLE OF INVENTION: EXONUCLEASE-DEFICIENT THERMOSTABLE I
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thomas Fitting
STREET: 12526 High Bluff Road, Suite 300
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92130
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/966,278
FILING DATE: 19921226
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: STE0133P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-792-3680
TELEFAX: 619-792-8677
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3499 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: 5'TIR
LOCATION: 1..223
NAME/KEY: CDS
LOCATION: 224..2551
FEATURE:
NAME/KEY: 3'TIR
LOCATION: 2552..3499
US-07-966-278-2

Query Match 79.0%; Score 15.8; DB 1; Length 3499;
Best Local Similarity 89.5%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAAGCGCCAGAGATGAT 19
||||| |||||||
DB 777 CAAGCGAGAGAGATGAT 795

RESULT 12
US-08-424-921-2
Sequence 2, Application US/08424921
Patent No. 5545552
GENERAL INFORMATION:
APPLICANT: Mathur, Eric A

;; TITLE OF INVENTION: PURIFIED THERMOSTABLE PYROCOCUS
;; TITLE OF INVENTION: FURIOSUS DNA POLYMERASE I
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Bingham & Filling
;; STREET: 12526 High Bluff Road, Suite 300
;; CITY: San Diego
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 92130
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/424,921
;; FILING DATE: 19-APR-1995
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/803,627
;; FILING DATE: 02-DEC-1991
;; APPLICATION NUMBER: US 07/620,568
;; FILING DATE: 03-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/657,073
;; FILING DATE: 19-FEB-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/776,552
;; FILING DATE: 15-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Filling, Thomas
;; REGISTRATION NUMBER: 34,163
;; REFERENCE/DOCKET NUMBER: STG0100P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619-792-3680
;; TELEFAX: 619-792-8477
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3499 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHEetical: NO
;; ANTI-SENSE: NO
;; FEATURE:
;; NAME/KEY: 5'UTR
;; LOCATION: 1..223
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 224..2551
;; FEATURE:
;; NAME/KEY: 3'UTR
;; LOCATION: 2552..3499
;; US-08-424-921-2
;;
;; Query Match 79.0%; Score 15.8; DB 1; Length 3499;
;; Best Local Similarity 89.5%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

;; TITLE OF INVENTION: furiosus DNA Polymerase I
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;; ADDRESSEE: Dunner, L.L.P.
;; STREET: 1300 I Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/556,355A
;; FILING DATE: 13-NOV-1995
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/424,921
;; FILING DATE: 19-APR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/803,627
;; FILING DATE: 02-DEC-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/779,846
;; FILING DATE: 21-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/657,073
;; FILING DATE: 19-FEB-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/620,568
;; FILING DATE: 03-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barker, M. Paul
;; REGISTRATION NUMBER: 32,013
;; REFERENCE/DOCKET NUMBER: 04121.0004-02
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)408-4000
;; TELEFAX: (202)408-4400
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3499 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-556-355A-2
;;
;; Query Match 79.0%; Score 15.8; DB 2; Length 3499;
;; Best Local Similarity 89.5%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGATGAT 19
||||| |||||||||
DB 777 CAAGCGAGAGAGATGAT 795

RESULT 14
US-07-803-627A-2
; Sequence 2, Application US/07803627A
; Patent No. 5948663
; GENERAL INFORMATION:
; APPLICANT: MATUR, Eric J.
; TITLE OF INVENTION: Purified Thermostable Pyrococcus
; TITLE OF INVENTION: furiosus DNA Polymerase I
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington

STATE: D.C.
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/803,627A
 FILING DATE: 02-DEC-1991
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/776,552
 FILING DATE: 14-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/657,073
 FILING DATE: 19-FEB-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/620,568
 FILING DATE: 03-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Barker, M. Paul
 REGISTRATION NUMBER: 32,013
 REFERENCE/DOCKET NUMBER: 04121.0004-00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)408-4000
 TELEFAX: (202)408-4400
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3499 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-07-803-627A-2

Query Match 79.0%; Score 15.8; DB 2; Length 3499;
 Best Local Similarity 89.5%; Pred. No. 60;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGAGATGAT 19
 ||||| |||||
 Db 777 CAAGCGAGAGAGATGAT 795

RESULT 15
 US-09-089-879-6
 ; Sequence 6, Application US/09089879
 ; Patent No. 611092
 ; GENERAL INFORMATION:
 ; APPLICANT: Williamson, Mark
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING DPT111 AND USES THEREOF
 ; FILE REFERENCE: 0734/078001
 ; CURRENT APPLICATION NUMBER: US/09/089,879
 ; CURRENT FILING DATE: 1998-06-03
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 1542
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)..(1542)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-089-879-6

Query Match 77.0%; Score 15.4; DB 3; Length 1542;
 Best Local Similarity 94.1%; Pred. No. 85;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AAGCGCAGAGAGATGA 18

|||||
 Db 483 AAGCGCAGAGAGAGA 499

Search completed: May 17, 2003, 03:41:13
 Job time : 30.1788 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 01:06:03 ; Search time 48.3444 Seconds
(without alignments)
533.646 Million cell updates/sec

Title: US-10-052-092-15

Perfect score: 20
Sequence: 1 caagcgccagagatgatg 20

Scoring table: IDENTITY_NDC

Gapop 10.0, Gapext 1.0

Searched: 810007 seqs, 644969091 residues

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|--------|---------------------|
| 1 | 20 | 100.0 | 9 | US-10-052-092-15 |
| 2 | 20 | 100.0 | 392 | US-10-052-092-4 |
| 3 | 20 | 100.0 | 1237 | US-10-052-092-3 |
| 4 | 20 | 100.0 | 1788 | US-10-052-092-24 |
| 5 | 20 | 100.0 | 1788 | US-09-853-033-1 |
| 6 | 20 | 100.0 | 2092 | US-10-052-092-6 |
| 7 | 20 | 100.0 | 2092 | US-09-952-680A-9 |
| 8 | 20 | 100.0 | 2106 | US-10-052-092-5 |
| 9 | 20 | 100.0 | 4963 | US-09-281-674-9 |
| 10 | 20 | 100.0 | 4963 | US-09-892-227-9 |
| 11 | 20 | 100.0 | 6450 | US-10-067-710-2 |
| 12 | 20 | 100.0 | 6450 | US-10-081-563-1 |
| 13 | 20 | 100.0 | 6450 | US-10-052-092-1 |
| 14 | 20 | 100.0 | 6450 | US-10-052-092-7 |
| 15 | 20 | 100.0 | 6610 | US-10-052-092-20 |
| 16 | 20 | 100.0 | 465237 | US-09-933-267A-1 |
| 17 | 16.4 | 82.0 | 1978 | US-10-052-092-23 |
| 18 | 15.8 | 79.0 | 426 | US-09-864-761-23769 |
| 19 | 15.8 | 79.0 | 534 | US-09-864-761-7038 |

| | | | | | | |
|----|------|------|--------|----|---------------------|--------------------|
| 20 | 15.8 | 79.0 | 2454 | 9 | US-10-060-036-176 | Sequence 176, App |
| 21 | 15.8 | 79.0 | 2454 | 10 | US-09-962-436-563 | Sequence 563, App |
| 22 | 15.8 | 79.0 | 2480 | 9 | US-10-060-036-4548 | Sequence 4548, App |
| 23 | 15.4 | 77.0 | 2000 | 9 | US-09-938-842A-2874 | Sequence 2874, App |
| 24 | 15.2 | 76.0 | 148 | 10 | US-09-764-877-681 | Sequence 681, App |
| 25 | 15.2 | 76.0 | 183 | 10 | US-09-920-300A-988 | Sequence 988, App |
| 26 | 15.2 | 76.0 | 183 | 12 | US-10-033-528-988 | Sequence 988, App |
| 27 | 15.2 | 76.0 | 210 | 10 | US-09-764-877-3406 | Sequence 3406, App |
| 28 | 15.2 | 76.0 | 395 | 9 | US-09-918-995-4369 | Sequence 4369, App |
| 29 | 15.2 | 76.0 | 419 | 9 | US-09-918-995-17697 | Sequence 17697, A |
| 30 | 15.2 | 76.0 | 427 | 9 | US-09-918-995-17725 | Sequence 17725, A |
| 31 | 15.2 | 76.0 | 791 | 10 | US-09-822-849A-578 | Sequence 578, App |
| 32 | 15.2 | 76.0 | 1002 | 10 | US-09-770-445-230 | Sequence 230, App |
| 33 | 15.2 | 76.0 | 1828 | 10 | US-09-917-800A-1643 | Sequence 1643, App |
| 34 | 15.2 | 76.0 | 145831 | 10 | US-09-969-708-79 | Sequence 79, App |
| 35 | 15.2 | 76.0 | 145831 | 10 | US-09-954-456-2116 | Sequence 2116, App |
| 36 | 15 | 75.0 | 442 | 9 | US-09-796-692-6887 | Sequence 6887, App |
| 37 | 15 | 75.0 | 442 | 9 | US-10-040-862-6887 | Sequence 6887, App |
| 38 | 15 | 75.0 | 485 | 9 | US-09-796-692-3867 | Sequence 3867, App |
| 39 | 15 | 75.0 | 485 | 9 | US-10-040-862-3867 | Sequence 3867, App |
| 40 | 15 | 75.0 | 504 | 9 | US-09-796-692-8497 | Sequence 8497, App |
| 41 | 15 | 75.0 | 504 | 9 | US-10-040-862-8497 | Sequence 8497, App |
| 42 | 15 | 75.0 | 1789 | 9 | US-09-924-340-109 | Sequence 109, App |
| 43 | 15 | 75.0 | 1789 | 9 | US-09-992-600A-109 | Sequence 109, App |
| 44 | 14.8 | 74.0 | 365 | 9 | US-09-736-457-391 | Sequence 391, App |
| 45 | 14.8 | 74.0 | 365 | 9 | US-09-902-941-391 | Sequence 391, App |

ALIGNMENTS

RESULT 1
US-10-052-092-15
; Sequence 15, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne
; APPLICANT: Allied, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therap
; FILE REFERENCE: P02102052
; CURRENT APPLICATION NUMBER: US/10/052,092
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-10-052-092-15
Query Match 100.0%; Score 20; DB 9; Length 20;
Best local similarity 100.0%; Pred. No. 0.82;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCGCCAGAGATGATG 20
|||||
Db 1 CAAGCGCCAGAGATGATG 20
RESULT 2
US-10-052-092-4
; Sequence 4, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fuqua, Suzanne

```

; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeuti
; FILE REFERENCE: P021020S2
; CURRENT APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Human
US-10-052-092-4

```

```

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 392;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 28 CAAGCGCCAGAGAGATGATG 47

```

```

RESULT 3
US-10-052-092-3
; Sequence 3, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fugua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeuti
; FILE REFERENCE: P021020S2
; CURRENT APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1237
; TYPE: DNA
; ORGANISM: Human
US-10-052-092-3

```

```

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 1237;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 390 CAAGCGCCAGAGAGATGATG 409

```

```

RESULT 4
US-10-052-092-24
; Sequence 24, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fugua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeuti
; FILE REFERENCE: P021020S2

```

```

; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Pig
US-10-052-092-24

```

```

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 1788;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 801 CAAGCGCCAGAGAGATGATG 820

```

```

RESULT 5
US-09-853-033-1
; Sequence 1, Application US/09853033
; Patent No. US20020100068A1
; GENERAL INFORMATION:
; APPLICANT: CHAMBER, PIERRE
; APPLICANT: METZGER, DANIEL
; TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION
; FILE REFERENCE: 065691/0222
; CURRENT APPLICATION NUMBER: US/09/853,033
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/12570
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1788)
US-09-853-033-1

```

```

Query Match
Best Local Similarity 100.0%; Score 20; DB 10; Length 1788;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAAGCGCCAGAGAGATGATG 20
Db 801 CAAGCGCCAGAGAGATGATG 820

```

```

RESULT 6
US-10-052-092-6
; Sequence 6, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fugua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therape
; FILE REFERENCE: P021020S2
; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018

```

PRIOR FILING DATE: 2001-07-09
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 6
 LENGTH: 2092
 TYPE: DNA
 ORGANISM: Human
 US-10-052-092-6

Query Match 100.0%; Score 20; DB 9; Length 2092;
 Best Local Similarity 100.0%; Pred. No. 0.93;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCCGAGAGATGATG 20
 ||||||||||||||||
 Db 1093 CAAGCCCGAGAGATGATG 1112

RESULT 7
 US-09-952-680A-9
 Sequence 9, Application US/09952680A
 Publication No. US20030087239A1
 GENERAL INFORMATION:
 APPLICANT: Stanton, Marty
 APPLICANT: Epstein, David
 APPLICANT: Hamaguchi, No. US20030087239A1
 TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
 FILE REFERENCE: 23239-501
 CURRENT APPLICATION NUMBER: US/09/952,680A
 CURRENT FILING DATE: 2001-09-13
 PRIOR APPLICATION NUMBER: 60/232,454
 PRIOR FILING DATE: 2000-09-13
 NUMBER OF SEQ ID NOS: 75
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 9
 LENGTH: 2092
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-952-680A-9

Query Match 100.0%; Score 20; DB 9; Length 2092;
 Best Local Similarity 100.0%; Pred. No. 0.93;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCCGAGAGATGATG 20
 ||||||||||||||||
 Db 1093 CAAGCCCGAGAGATGATG 1112

RESULT 8
 US-10-052-092-5
 Sequence 5, Application US/10052092
 Publication No. US20030027778A1
 GENERAL INFORMATION:
 APPLICANT: Fugua, Suzanne
 APPLICANT: Allred, D.
 APPLICANT: Hopp, Torsten A.
 APPLICANT: O'Connell, Peter
 TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeutics
 FILE REFERENCE: P02102052
 CURRENT APPLICATION NUMBER: US/10/052,092
 CURRENT FILING DATE: 2002-01-18
 PRIOR APPLICATION NUMBER: US 60/262,990
 PRIOR FILING DATE: 2001-01-19
 PRIOR APPLICATION NUMBER: US 60/304,018
 PRIOR FILING DATE: 2001-07-09
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 5
 LENGTH: 2106
 TYPE: DNA
 ORGANISM: Human
 US-10-052-092-5

Query Match 100.0%; Score 20; DB 9; Length 2106;
 Best Local Similarity 100.0%; Pred. No. 0.93;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCCGAGAGATGATG 20
 ||||||||||||||||
 Db 801 CAAGCCCGAGAGATGATG 820

RESULT 9
 US-09-281-674-9
 Sequence 9, Application US/09281674
 Patent No. US20020077307A1
 GENERAL INFORMATION:
 APPLICANT: Gossen, Manfred
 APPLICANT: Bujard, Hermann
 APPLICANT: Salfield, Jochen
 APPLICANT: Voss, Jeffrey
 TITLE OF INVENTION: Methods for Regulating Gene Expression
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lohive & Cockfield
 STREET: 60 State Street, Suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/281,674
 FILING DATE: 30-Mar-1999
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/479,306
 FILING DATE: <Unknown>
 APPLICATION NUMBER: 08/260,452
 FILING DATE: 14-JUN-1994
 APPLICATION NUMBER: 08/076,327
 FILING DATE: 14-JUN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Giulio A. DeConti, Jr.
 REGISTRATION NUMBER: 31,503
 REFERENCE/DOCKET NUMBER: Brl-013CP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4963 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Human cytomegalovirus
 IMMEDIATE SOURCE:
 CLONE: PUHD BGR4
 SEQUENCE DESCRIPTION: SEQ ID NO: 9:
 US-09-281-674-9

Query Match 100.0%; Score 20; DB 10; Length 4963;
 Best Local Similarity 100.0%; Pred. No. 0.95;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCCGAGAGATGATG 20
 ||||||||||||||||
 Db 1266 CAAGCCCGAGAGATGATG 1285

RESULT 10
US-09-892-227-9
Sequence 9, Application US/09892227
Patent No. US20020152487A1
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
Bujard, Hermann
Salzfeld, Jochen
Voss, Jeffrey
TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled Transcription
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Iahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/892,227
FILING DATE: 25-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/163,269
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-Jun-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4963 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
IMMEDIATE SOURCE:
CLONE: pUHD BGR4
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-892-227-9
Query Match 100.0%; Score 20; DB 10; Length 4963;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGGCCAGAGAGATGATG 20
DB 1266 CAAGGCCAGAGAGATGATG 1285

RESULT 11
US-10-096-710-2
Sequence 2, Application US/10096710
Patent No. US20020164581A1
GENERAL INFORMATION:
APPLICANT: Chang, Chawmsang
TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS THAT
TITLE OF INVENTION: DO NOT ATTENUATE THE PROTECTIVE EFFECTS OF ESTROGEN
FILE REFERENCE: 21108.001002

CURRENT APPLICATION NUMBER: US/10/096,710
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/275,229
PRIOR FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 2
LENGTH: 6450
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:/note =
OTHER INFORMATION: synthetic construct
US-10-096-710-2
Query Match 100.0%; Score 20; DB 9; Length 6450;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGGCCAGAGAGATGATG 20
DB 1161 CAAGGCCAGAGAGATGATG 1180

RESULT 12
US-10-081-563-1
Sequence 1, Application US/10081563
Publication No. US20020187495A1
GENERAL INFORMATION:
APPLICANT: Wake Forest University
APPLICANT: Harrington, David M.
APPLICANT: Howard, Timothy D.
APPLICANT: Hawkins, Gregory A.
TITLE OF INVENTION: GENETIC POLYMORPHISMS OF ESTROGEN RECEPTOR ALPHA ASSOCIATED WITH
FILE REFERENCE: 9151-15
CURRENT APPLICATION NUMBER: US/10/081,563
CURRENT FILING DATE: 2002-02-22
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 1
LENGTH: 6450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (361)..(2148)
OTHER INFORMATION:
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: X03635
DATABASE ENTRY DATE: 1993-09-12
RELEVANT RESIDUES: (1)..(6450)
US-10-081-563-1
Query Match 100.0%; Score 20; DB 9; Length 6450;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGGCCAGAGAGATGATG 20
DB 1161 CAAGGCCAGAGAGATGATG 1180

RESULT 13
US-10-052-092-1
Sequence 1, Application US/10052092
Publication No. US20030027778A1
GENERAL INFORMATION:
APPLICANT: Fuqua, Suzanne
APPLICANT: Allied, D.
APPLICANT: Hopp, Torsten A.
APPLICANT: O'Connell, Peter

```

; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapent
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6450
; TYPE: DNA
; ORGANISM: Human
US-10-052-092-1

```

```

Query Match          100.0%; Score 20; DB 9; Length 6450;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 CAAGGCCGAGAGATGATG 20
        ||||||||||||||||
Db      1161 CAAGGCCGAGAGATGATG 1180

```

```

RESULT 14
US-10-052-092-7
; Sequence 7, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fugua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapent
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 6450
; TYPE: DNA
; ORGANISM: Human
US-10-052-092-7

```

```

Query Match          100.0%; Score 20; DB 9; Length 6450;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 CAAGGCCGAGAGATGATG 20
        ||||||||||||||||
Db      1161 CAAGGCCGAGAGATGATG 1180

```

```

RESULT 15
US-10-052-092-20
; Sequence 20, Application US/10052092
; Publication No. US20030027778A1
; GENERAL INFORMATION:
; APPLICANT: Fugua, Suzanne
; APPLICANT: Allred, D.
; APPLICANT: Hopp, Torsten A.
; APPLICANT: O'Connell, Peter
; TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapent
; FILE REFERENCE: P02102US2
; CURRENT APPLICATION NUMBER: US/10/052,092
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,990

```

```

; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/304,018
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 6610
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector pCI-n GLL-HEGO
US-10-052-092-20

```

```

Query Match          100.0%; Score 20; DB 9; Length 6610;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 CAAGGCCGAGAGATGATG 20
        ||||||||||||||||
Db      2685 CAAGGCCGAGAGATGATG 2704

```

```

Search completed: May 17, 2003, 11:07:21
Job time : 56.3444 secs

```

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|-------------|
| c 1 | 20 | 100.0 | 653 | 9 | A1073549 | 0v45e08_x |
| | 18.4 | 99.0 | 641 | 10 | AP076636 | ESP388744 |
| c 3 | 18 | 90.0 | 642 | 12 | BP058827 | 7K13b08-7 |
| c 4 | 17.4 | 87.0 | 687 | 10 | BES13707 | 60J318478 |
| c 5 | 17.4 | 87.0 | 771 | 10 | AQ249357 | T10H18-SF |
| c 6 | 16.8 | 84.0 | 267 | 17 | BE154100 | PM1-HT034 |

RESULT 1
 LOCUS AI073549/c
 DEFINITION ov45e08.x1 Soares_testis_MRT Homo sapiens cDNA clone IMAGE:1640294
 3', similar to gb:MI2674 ESTROGEN RECEPTOR (HUMAN);, mRNA sequence.
 ACCESSION AI073549
 VERSION AI073549.1 GI:3400193
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 653)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bbrp/image/image.html
 Seq primer: -40m3 fwd. RT from Amersham
 High quality sequence stop: 483.
 Location/Qualifiers

source

1. 653

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="1640294"

/clone_lib="Soares_testis_NHT"

/sex="male"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was prepared from mRNA obtained from Clontech Laboratories

, Inc., and primed with a Not I - oligo(dT) primer [5'

TGTACCATCTGAGTGGAGCGCGCCGCAATTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization to Cols, and was

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 161 a 169 c 142 g 181 t

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 653;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCCGAGAGATGATG 20

DB 612 CAAGCCGAGAGATGATG 593

RESULT 2

LOCUS AM976636 641 bp mRNA linear EST 02-JUN-2000

DEFINITION EST388745 MAGE resequences, MAGN Homo sapiens cDNA, mRNA sequence.

ACCESSION AM976636

VERSION AM976636.1 GI:8167867

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 641)

Hegde, P., Qi, R., Aernathy, K., Dharap, S., Gaspar, R., Gay, C., Holt

, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J., and

Quackenbush, J.

Assessment of gene expression patterns in a model of colon tumor

metastasis using a 19,200 element cDNA microarray

Unpublished (2000)

Contact: John Quackenbush

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 3528

Fax: 301 838 0208

Email: johnq@tigr.org

Plate: 363

Seq primer: Forward.

Location/Qualifiers

1. 641

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="MAGE resequences, MAGN"

/note="Vector: pBluescriptSK"

BASE COUNT 159 a 175 c 193 g 112 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 641;

Best Local Similarity 95.0%; Pred. No. 1.9e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAAGCCGAGAGATGATG 20

DB 597 CAAGCCGAGAGATGATG 616

RESULT 3

BF058527 642 bp mRNA linear EST 16-OCT-2000

LOCUS

7K31H08.x1 NCI-CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3477350 3'

similar to SW:ESRL_HUMAN P03372 ESTROGEN RECEPTOR 1; mRNA sequence.

ACCESSION

BF058527

VERSION

BF058527.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 642)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaaps-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskalko, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: K. Bento

Soares, Ph.D. CDNA Library Arrayed by: Christa Prange, The

I.M.A.G.E. Consortium DNA Sequencing by: Washington University

Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/ILMT, send email to:

info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 492.

Location/Qualifiers

1. 642

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="3477350"

/clone_lib="NCI-CGAP_Ov18"

/tissue_type="fibrotheoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: ovary; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTACCATCTGAGTGGAGCGCGCCGCAATTTTTTTTTTTT 3'];

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 158 a 182 c 196 g 106 t

ORIGIN

Query Match 90.0%; Score 18; DB 12; Length 642;

Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCCGAGAGATGATG 18

DB 609 CAAGCCGAGAGATGATG 626

RESULT 4

BE513707 687 bp mRNA linear EST 07-AUG-2000

LOCUS BE513707

DEFINITION BE513707

MAGE resequence.

ACCESSION BE513707

VERSION BE513707.1

GI:9720919

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 687)

NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@b-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 DNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNU at: image.lnl.gov.
 Plate: ILICM328 row: e column: 19
 High quality sequence start: 6
 High quality sequence stop: 441.

FEATURES

source

1. 687

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="3634170"

/clone_lib="NH_MGC-8"

/issue_type="Burkitt lymphoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph. Vector: pDBT; Site.1: XhoI; Site.2:

EcotRI; CDNA made by oligo-dT priming. Directionally

cloned into EcotRI/XhoI sites using the following 5'

adaptor: GGCACGAG(g). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-CDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 184 a 176 c 207 g 120 t

ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 687;

Best Local Similarity 94.7%; Pred. No. 6e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AACGCCAGAGAGATGATG 20

DB 663 AACGCCAGAGAGATGAG 681

RESULT 5

LOCUS A0249357 771 bp DNA linear GSS 06-OCT-1998

DEFINITION T10H18-Sp6 TAMU Arabidopsis thaliana genomic clone T10H18, DNA

sequence.

ACCESSION A0249357

VERSION A0249357.1 GI:3699440

KEYWORDS GSS.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 771)

Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and

Ecker, J.

BAC End Sequences at ATGC

Unpublished (1997)

Contact: Ecker, J.

Arabidopsis thaliana Genome Center

University of Pennsylvania

Dept. of Biology, University of Pennsylvania, Philadelphia, PA

19104

Tel: 215-898-9384

Fax: 215-898-8780

Email: jecker@atgenom.bio.upenn.edu

Seq primer: Sp6

Class: BAC ends.

Location/Qualifiers

1. 771

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="T10H18"
 /clone_lib="TAMU"
 /sex="thermaphrodite"
 /note="Vector: DeloBAC1; Site.1: HindIII; Site.2: HindIII
 ; Produced by Rod Wing"

BASE COUNT 281 a 154 c 108 g 201 t 27 others

ORIGIN

Query Match 87.0%; Score 17.4; DB 17; Length 771;

Best Local Similarity 94.7%; Pred. No. 6.5e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAGCCAGAGAGATGAT 19

DB 73 CAAGCCAGAGAGATGAT 55

RESULT 6

LOCUS BE154100/c 267 bp mRNA linear EST 21-JUN-2000

DEFINITION PM1-HT0340-091199-001-808 HT0340 Homo sapiens CDNA, mRNA sequence.

ACCESSION BE154100

VERSION BE154100.1 GI:8616821

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 267)

Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M. R.,

Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,

Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,

Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare

, M. J., Soares, P., Brentani, R. R., Reis, L. F., de Souza, S. J. and

Simpson, A. V.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2-PM1-HT0340-091>)

199-001-008613-1999-11-09&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 22

High quality sequence stop: 267.

Location/Qualifiers

1. 267

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0340"

/dev_stage="Adult"

/note="Organ: head/neck; Vector: puc18; Site.1: SmaI;

Site.2: SmaI; A mini-library was made by cloning products

derived from ORSETES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 67 a 63 c 77 g 60 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 267;

Best Local Similarity 90.0%; Pred. No. 6.9e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGATGATG 20
 DB 146 CAAGCGCCAGAGATGATG 127

RESULT 7
 AM354950/c 319 bp mRNA linear EST 25-APR-2001

LOCUS 37558 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
 AM354950
 ACCESSION AM354950
 VERSION AM354950.1 GI:6853940
 KEYWORDS EST.
 SOURCE COW
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 319)
 Smith, R.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
 Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett,
 G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A., Chitko-McKown, C.G.,
 Perle, G., Holt, L.I., Karamycheva, S., Liang, F., Quackenbush, J. and
 Keeler, J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)

TITLE
 JOURNAL MEDLINE
 21180013

COMMENT
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemail@marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -mismore 20
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCATGATGACAT
 BACKWARD: GTTTCGACATGACGACG
 Plate: 22 row: J column: 16
 Seq primer: ATTGAGTGACATGATG.
 Location/Qualifiers
 1..319
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 1BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
 library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."

BASE COUNT 79 a 109 g 54 t
 ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 319;
 Best Local Similarity 90.0%; Pred. No. 7.6e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGATGATG 20
 DB 60 CAAGCGCCAGAGATGATG 41

RESULT 8
 BM660907/c 327 bp mRNA linear EST 27-FEB-2002

LOCUS 952042H04.xl 952 - BMS tissue from Walbot Lab (reduced rRNA) Zea
 DEFINITION
 mays cDNA, mRNA sequence.
 BM660907
 ACCESSION BM660907
 VERSION BM660907.1 GI:18964264
 KEYWORDS EST.
 SOURCE Zea mays.

ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 Clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 327)
 Walbot, V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 952042 row: H column: 04.
 Location/Qualifiers
 1..327
 /organism="Zea mays"
 /cultivar="BMS (Black Mexican Sweet)"
 /db_xref="taxon:4577"
 /clone_lib="952 - BMS tissue from Walbot Lab (reduced rRNA)"
 /tissue_type="suspension culture"
 /dev_stage="mixed logarithmic and stationary growth
 phases"
 /lab_host="DH10B"
 /note="Vector: pUC19; Site_1: EcoRI; Site_2: EcoRI. The
 library was prepared by George Rudenko using poly (A)
 selected RNA and Universal Riboclone cDNA Synthesis System
 (Promega). cDNA was synthesized using both random and
 oligo(dT) primers in separate reactions and equipped with
 EcoRI adaptors. Library was size-fractionated on agarose
 gels (for insert size >400bp) and non-directionally cloned
 into EcoRI-digested pUC19 vector. Blue/white selection on
 carbenicillin-containing plates was used to recover
 positive clones."

BASE COUNT 86 a 88 c 65 g 86 t 2 others
 ORIGIN

Query Match 84.0%; Score 16.8; DB 13; Length 327;
 Best Local Similarity 90.0%; Pred. No. 7.7e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAGCGCCAGAGATGATG 20
 DB 302 CAAGCGCCAGAGATGATG 283

RESULT 9
 BE592363 370 bp mRNA linear EST 18-AUG-2000

LOCUS BE592363
 DEFINITION mRNA sequence.
 WS1_93_H01.g1_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA.
 BE592363
 ACCESSION BE592363
 VERSION BE592363.1 GI:9847436
 KEYWORDS EST.
 SOURCE sorghum.
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 Clade; Panicoidae; Andropogoneae; Sorghum.
 1 (bases 1 to 370)
 Cordonnier-Pratt, M.-M., Gingle, A., Marasala, C., Sudman, M. and Pratt,
 L.H.
 An EST database from Sorghum: water-stressed plants
 Unpublished (2000)
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860

Fax: 706 583 0210
 Email: mmpatt@uga.edu
 Sequences have been trimmed to exclude polyA, vector and regions
 below phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: PolYTmX
 High quality sequence start: 6
 High quality sequence stop: 352
 POLYA-No.

FEATURES

SOURCE

Location/Qualifiers

1..370
 /organism="Sorghum bicolor"
 /db_xref="taxon:4558"
 /clone_lib="Water-stressed 1 (WS1)"
 /note="Organ: Mix of 5-week old plants on days 7 & 8 after
 water was withheld; Vector: lambda zap; Site_1: XhoI;
 Site_2: EcoRI; The library was made from polyA RNA in the
 cloning vector lambda ZAP II. Clones to be sequenced were
 prepared by mass excision."

BASE COUNT

100 a 76 c 103 g 90 t 1 others

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 370;
 Best Local Similarity 90.0%; Pred. No. 8.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAGCGCCGAGAGATGATG 20

Db 40 CAGCGAAGAGAGATGATG 59

RESULT 10

BF042797/c

LOCUS BF042797 386 bp mRNA linear EST 10-OCT-2000
 DEFINITION BF250002A10F10 Soares normalized bovine placenta Bos taurus cDNA

Accession BF042797

Version BF042797

Keywords BF042797.1 GI:10759852

Source EST.

Organism COW.

Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 386)

AUTHORS Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,T., Liu,L. and Larson

TITLE 'U.H.

JOURNAL Bovine ESTS

COMMENT Unpublished (2000)

W. M. Keck Center for Comparative and Functional Genomics

University of Illinois at Urbana-Champaign

340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL

61801, USA

Tel: 217 333 5998

Fax: 217 244 5617

Email: h-lewin@uiuc.edu

Funding for cattle EST sequencing was provided by the USDA National
 Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
 to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
 from Washington University Genome Center. Vector Trimmi g:

Cross match from Washington University Genome Center PHRAP suite.

This sequence is vector free and at least 200 bp in length. REPEAT

IN THE SEQUENCE low_complexity STRAND (+) ELEMENT GC-rich LOCATION

[30,81].

PCR primers

FORWARD: TAATAGCATCATATAGG

BACKWARD: ATTAACCTCATAAG

Insert Length: 386 Std Error: 0.00

Plate: BP250002A10 row: F column: 10

Seq primer: ACGGATACCAATTTCACACAGA

High quality sequence stop: 386.

Location/Qualifiers

FEATURES

SOURCE

1..386

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone="BP250002A10F10"

/clone_lib="Soares normalized bovine placenta"

/sex="female"

/lab_host="DH10B"

/note="Organ: placenta; Vector: pUT3pac; Site_1: EcoRI;
 Site_2: NotI; The cDNA library was contributed by the
 Soares laboratory and it was constructed and normalized
 as described by Bonaldo, M.F., Lennon, G. and Soares,
 M.B. (1996), Genome Research 6(9): 791-806."

BASE COUNT 83 a 103 c 141 g 59 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 386;
 Best Local Similarity 90.0%; Pred. No. 8.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAGCGCCGAGAGATGATG 20

Db 220 CGAGCGCCGAGAGATGATG 201

RESULT 11

BH552907/c

LOCUS BH552907 422 bp DNA linear GSS 14-DEC-2001
 DEFINITION BOHQ61R BOHQ Brassica oleracea genomic clone BOHQ61, DNA

Accession BH552907

Version BH552907

Keywords BH552907.1 GI:17804687

Source GSS.

Organism Brassica oleracea.

Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 422)

AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.

TITLE Whole genome shotgun sequencing of Brassica oleracea

JOURNAL Unpublished (2001)

COMMENT Other GSSs: BOHQ61R

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES

Location/Qualifiers

1..422

/organism="Brassica oleracea"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOHQ61"

/clone_lib="BOHQ"

/note="Vector: pHS01; Site_1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHS01 using BstXI linkers"

BASE COUNT 120 a 108 c 94 g 100 t

ORIGIN

Query Match 84.0%; Score 16.8; DB 17; Length 422;
 Best Local Similarity 90.0%; Pred. No. 8.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAGCGCCGAGAGATGATG 20

Db 62 CAAGCGCCGAGAGATGATG 43

RESULT 12

| | | | | |
|------------|---|--|--|--|
| TITLE | Unpublished (2001) Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.go.jp, url:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh ,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res., 10 (10), 1617-1630 (2000) Wegli,K., Fujiwara,S., Inoue,K., Togawa,Y., Iwawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Komno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara ,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res., 11 (2), 281-289 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues. | | | |
| FEATURES | Location/Qualifiers | | | |
| SOURCE | 1.438 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="F930004E21" /clone_id="RIKEN full-length enriched, adult inner ear" /tissue_type="inner ear" /dev_stage="adult" /note="Pooled tissues: (tissue_type=cerebellum, dev_stage=16 days neonate, sex=mixed), (tissue_type=cerebellum, dev_stage=0 day neonate, sex=mixed), (tissue_type=hippocampus, dev_stage=adult, sex=male), (tissue_type=whole body, dev_stage=9 days embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days embryo, sex=mixed)" | | | |
| BASE COUNT | 100 a 119 c 143 g 76 t | | | |
| ORIGIN | Query Match 84.0%; Score 16.8; DB 10; Length 438; Best Local Similarity 90.0%; Pred. No. 9, 1e+02; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0; QY 1 CAAGCGCCAGAGAGATGATG 20 11111111111111111111 Db 160 CAAGCGCCAGAGAGATGCTG 179 | | | |
| RESULT 14 | BB847728 | | | |
| LOCUS | BB847728 439 bp mRNA linear EST 26-NOV-2001 | | | |
| DEFINITION | BB847728 RIKEN full-length enriched, adult male kidney Mus musculus | | | |
| ACCESSION | CDNA clone F530101F07 5', mRNA sequence. | | | |
| VERSION | BB847728 | | | |
| KEYWORDS | BB847728.1 GI:17086103 EST. | | | |

Mon May 19 10:04:46 2003

us-10-052-092-15.rst

Page 8

Job time : 937.993 secs
